

Type	L #	Hits	Search Text	DBs	Time Stamp	Comments	Error Definition	Errors
1 BRS	L28	122	staphylococcus adj aureus	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:29			0
2 BRS	L29	0	s. adj aureus	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:29			0
3 BRS	L30	2	dnai adj polypeptide	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:30			0
4 BRS	L31	24	dnai	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:30			0
5 BRS	L33	0	28 same (30 or 31)	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:33			0
6 BRS	L34	2	31 same binding	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:33			0
7 BRS	L35	1	inhibitor same 34	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:34			0
8 BRS	L36	46988	bacteri\$2 same infection	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:35			0
9 BRS	L37	2	36 same 30	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:38			0
10 BRS	L38	1	31 same inhibitor	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:39			0

THIS PAGE BLANK (USPTO)

	Type	L #	Hits	Search Text	DBs	Time Stamp	Comments	Error Definition	Errors
11	BRS	L32	1	bacteriophage adj "77" adj orf adj "104"	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:40			0
12	BRS	L39	1	(32 or 34) same inhibitor	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:40			0
13	BRS	L40	7	dnac adj helicase	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:40			0
14	BRS	L41	1	31 same 40	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:42			0
15	BRS	L42	3	31 same (fragment or variant)	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:43			0
16	BRS	L43	0	36 same 42	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:43			0
17	BRS	L44	10	pelletier adj jerry.in.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:44			0
18	BRS	L46	12	gros adj philippe.in.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:44			0
19	BRS	L47	1	dubow adj mike.in.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:45			0
20	BRS	L48	1	(44 or 46 or 47) and 31	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/16 16:45			0

THIS PAGE BLANK (USPTO)

	Type	L #	Hits	Search Text	DBs	Time Stamp	Comments	Error Definition	Errors
1	BRS	L1	24	dnai	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 14:58			0
2	BRS	L2	37748	bacteri\$2 same inhibits\$3	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 14:58			0
3	BRS	L3	2	1 same 2	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 14:59			0
4	BRS	L4	1	bacteriophage adj "77" adj orf adj "104"	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 15:03			0
5	BRS	L5	2	3 same 1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 15:00			0
6	BRS	L6	1	4 same 1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 15:00			0
7	BRS	L7	24094	bacteriophage	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 15:01			0
8	BRS	L8	434	2 same 7	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 15:01			0
9	BRS	L9	5	bacteriophage adj inhibits\$3 adj protein	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 15:03			0
10	BRS	L10	0	1 same 9	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/01/17 15:04			0

THIS PAGE BLANK (USPTO)

=> d his

(FILE 'HOME' ENTERED AT 16:50:41 ON 16 JAN 2004)

FILE 'MEDLINE, CAPLUS, BIOSIS, EMBASE, SCISEARCH, AGRICOLA'
ENTERED AT

16:51:01 ON 16 JAN 2004

L1 223 S STAPHYLOCOCCUS AUREUS
L2 53545 S S. AUREUS
L3 53722 S L1 OR L2
L4 148 S DNAI
L5 3 S L3 (P) L4
L6 3 S INHIBIT? (P) L5
L7 2 DUPLICATE REMOVE L6 (1 DUPLICATE REMOVED)
L8 375307 S BACTERI? (P) INFECTION
L9 3 S L4 (P) L8
L10 3 DUPLICATE REMOVE L9 (0 DUPLICATES REMOVED)
L11 1 S L10 NOT L7
L12 2 S BACTERIOPHAGE 77 ORF 104
L13 2 S L12 (P) L4
L14 0 S L13 NOT L9
L15 2 S L4 (P) INHIBITOR
L16 2 S L15 NOT L9
L17 1 DUPLICATE REMOVE L16 (1 DUPLICATE REMOVED)
L18 16 S L4 (P) (FRAGMENT OR VARIANT)
L19 3 S L18 (P) L3
L20 1 S L19 NOT L9
L21 0 S PELLETIER J?/AU
L22 3615 S PELLETIER J?/AU
L23 2074 S GROS P?/AU
L24 393 S DUBOW M?/AU
L25 6047 S (L22 OR L23 OR L24)
L26 3 S L25 AND L4
L27 2 DUPLICATE REMOVE L26 (1 DUPLICATE REMOVED)
L28 0 S L27 NOT L9

=> log y

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 23.912 Seconds

(without alignments) 1086.625 Million cell updates/sec

Title: US-09-689-952-16

Perfect score: 851
Sequence: 1 AADICTAITNGEQVKGTL.....ERVKSISPTPYFLSGENFRNN 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: A.Geneseq.19jun03.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	851	100.0	164	22	AA047315
2	851	100.0	250	22	AA047316
3	851	100.0	313	22	AA047317
4	782	91.9	299	22	AA082049
5	782	91.9	299	22	AA082337
6	782	91.9	307	23	ABP38947
7	418.5	49.2	307	23	AB049355
8	393	46.2	92	23	ABP09981
9	351	41.2	300	23	ABP27829

10	346	40.7	298	24	AB02187
11	332.5	39.1	300	23	ABP27828
12	330	38.8	293	23	AB054065
13	104.5	12.3	248	22	AA099665
14	104	12.0	386	22	AB052088
15	102.5	12.0	304	23	ABP40308
16	101	11.9	294	23	AB054759
17	100	11.8	291	23	AB053757
18	99	11.6	855	22	AA096545
19	98.5	11.6	281	22	AB055583
20	98.5	11.6	281	22	AB018846
21	98	11.5	467	22	AA034802
22	98	11.5	479	22	AB018303
23	97.5	11.5	798	22	AA096715
24	97	11.4	435	22	AB025176
25	95.5	11.2	340	21	AA052000
26	95.5	11.2	340	21	AA051629
27	95.5	11.2	441	19	AA080725
28	95.5	11.2	442	22	AA037959
29	95.5	11.2	453	24	AB027771
30	95.5	11.2	784	22	AB030180
31	95	11.2	383	24	ABP78794
32	94	11.0	363	22	AA096616
33	94	11.0	451	23	AB049805
34	94	11.0	453	23	ABP30561
35	94	11.0	484	23	ABP27967
36	93.5	11.0	892	24	AB074125
37	93	10.9	500	23	AB065636
38	92.5	10.9	574	22	AA017209
39	92.5	10.9	893	24	AB074127
40	92.5	10.9	893	24	AB019332
41	92.5	10.9	955	24	AB074126
42	92	10.8	264	18	AA021778
43	92	10.8	433	24	ABP77452
44	92	10.8	439	22	AB061378
45	92	10.8	440	17	AA094600

ALIGNMENTS

RESULT 1
ID AAB47315
AAB47315 standard; Protein; 164 AA.
XX AAB47315;
AC AAB47315;
XX 29-AUG-2001 (first entry)
DT
XX Amino acids 150-313 of S. aureus DnaI.
DE
XX DnaI; S. aureus; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
KW screening assay.
XX Staphylococcus aureus.
OS
XX WO200146383-A2.
XX 28-JUN-2001.
XX 21-DEC-2000; 2000WO-US35180.
XX 22-DEC-1999; 99US-0470512.
XX 12-OCT-2000; 2000US-0689952.
XX (PHAG-) PHAGETECH INC.
XX (WILL) WILLIAMS K M.
XX Pelletier J, Gros P, Dubow M;
XX WPI, 2001-418052/44.
XX N-PSDB; AAC86104.

PT Novel DnaI polypeptides useful for treating and diagnosing microbial,
PT preferably bacterial, diseases such as those caused by *Staphylococcus*
XX aureus -
XX
XX Claim 41; Fig 15; 107pp; English.
XX
CC This sequence shows a DnaI polypeptide derived from *S. aureus*. *S.*
CC aureus DnaI interacts with a growth-inhibitory bacteriophage 77 ORF
CC 104 gene product, to form the basis of a screening assay. DnaI
CC polypeptides and polynucleotides are useful for treating microbial,
CC preferably bacterial, especially *Staphylococcus*, infections. DnaI
CC polypeptides and polynucleotides are useful for biological, diagnostic,
CC prophylactic, clinical and therapeutic use, and as components in
CC databases useful for search analyses as well as in sequence analysis
CC algorithms.
XX
XX Sequence 164 AA;
SQ
Query Match 100.0%; Score 851; DB 22; Length 164;
Best Local Similarity 100.0%; Pred. No. 9.6e-87;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AADDICTAINTNGEQVGLVYGPFGTGSFLLGAINQLKSKVRSITIIYLPFIRTLKG 60
DB 1 AADDICTAINTNGEQVGLVYGPFGTGSFLLGAINQLKSKVRSITIIYLPFIRTLKG 60
QY 61 GFDGSGFEKKLHVRANIIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNFD 120
DB 61 GFDGSGFEKKLHVRANIIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNFD 120
QY 121 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFSGENFRNN 164
DB 121 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFSGENFRNN 164
RESULT 2
AAB47316
ID AAB47316 standard; Protein; 250 AA.
XX
XX AAB47316;
AC
XX
XX 29-AUG-2001 (first entry)
DT
XX
XX Amino acids 64-313 of *S. aureus* DnaI.
DE
XX
XX DnaI; *S. aureus*; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
KM screening assay.
XX
XX *Staphylococcus aureus*.
OS
XX
XX WO200146383-A2.
PN
XX
XX 28-JUN-2001.
PD
XX
XX 21-DEC-2000; 2000WO-US35180.
PF
XX
XX 22-DEC-1999; 99US-0470512.
PR
XX
XX 12-OCT-2000; 2000US-0689952.
PR
XX
XX (PHAG-) PHAGETECH INC.
PA (WILL/) WILLIAMS K M.
PA
XX
XX Pelletier J, Gros P, Dubow M;
PI
XX
XX WPI; 2001-418052/44.
DR
XX
XX Novel DnaI polypeptides useful for treating and diagnosing microbial,
PT preferably bacterial, diseases such as those caused by *Staphylococcus*
PT aureus -
XX
XX Disclosure; Fig 15; 107pp; English.
PS
XX
XX This sequence shows a DnaI polypeptide derived from *S. aureus*. *S.*

CC aureus DnaI interacts with a growth-inhibitory bacteriophage 77 ORF
CC 104 gene product, to form the basis of a screening assay. DnaI
CC polypeptides and polynucleotides are useful for treating microbial,
CC preferably bacterial, especially *Staphylococcus*, infections. DnaI
CC polypeptides and polynucleotides are useful for biological, diagnostic,
CC prophylactic, clinical and therapeutic use, and as components in
CC databases useful for search analyses as well as in sequence analysis
CC algorithms.
XX
XX Sequence 250 AA;
SQ
Query Match 100.0%; Score 851; DB 22; Length 250;
Best Local Similarity 100.0%; Pred. No. 1.7e-86;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AADDICTAINTNGEQVGLVYGPFGTGSFLLGAINQLKSKVRSITIIYLPFIRTLKG 60
DB 87 AADDICTAINTNGEQVGLVYGPFGTGSFLLGAINQLKSKVRSITIIYLPFIRTLKG 146
QY 61 GFDGSGFEKKLHVRANIIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNFD 120
DB 147 GFDGSGFEKKLHVRANIIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNFD 206
QY 121 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFSGENFRNN 164
DB 207 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFSGENFRNN 250
RESULT 3
AAB47317
ID AAB47317 standard; Protein; 313 AA.
XX
XX AAB47317;
AC
XX
XX 29-AUG-2001 (first entry)
DT
XX
XX *S. aureus* DnaI.
DE
XX
XX DnaI; *S. aureus*; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
KM screening assay.
XX
XX *Staphylococcus aureus*.
OS
XX
XX WO200146383-A2.
PN
XX
XX 28-JUN-2001.
PD
XX
XX 21-DEC-2000; 2000WO-US35180.
PF
XX
XX 22-DEC-1999; 99US-0470512.
PR
XX
XX 12-OCT-2000; 2000US-0689952.
PR
XX
XX (PHAG-) PHAGETECH INC.
PA (WILL/) WILLIAMS K M.
PA
XX
XX Pelletier J, Gros P, Dubow M;
PI
XX
XX WPI; 2001-418052/44.
DR
XX
XX N-PSDB; AAC86105.
DR
XX
XX Novel DnaI polypeptides useful for treating and diagnosing microbial,
PT preferably bacterial, diseases such as those caused by *Staphylococcus*
PT aureus -
XX
XX Disclosure; Fig 1; 107pp; English.
PS
XX
XX This sequence shows DnaI derived from *S. aureus*. *S. aureus* DnaI
CC interacts with a growth-inhibitory bacteriophage 77 ORF 104
CC gene product, to form the basis of a screening assay. DnaI
CC polypeptides and polynucleotides are useful for treating microbial,
CC preferably bacterial, especially *Staphylococcus*, infections. DnaI
CC polypeptides and polynucleotides are useful for biological, diagnostic,
CC prophylactic, clinical and therapeutic use, and as components in

CC databases useful for search analyses as well as in sequence analysis
CC algorithms.

XX Sequence 313 AA;

Query Match 100.0%; Score 851; DB 22; Length 313;
Best Local Similarity 100.0%; Pred. No. 2,3e-86;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AADDTCTAATNGEQYKGLYLPGPFGTGSFIIIGALANOLKSKVNSTIYVLEPIRTLKG 60
DB 150 AADDTCTAATNGEQYKGLYLPGPFGTGSFIIIGALANOLKSKVNSTIYVLEPIRTLKG 209
QY 61 GFKDSFEKKLRRVREANITLMDIGAEVTPWVDEVIIGPLHYRMVHELPTFSSNPD 120
DB 210 GFKDSFEKKLRRVREANITLMDIGAEVTPWVDEVIIGPLHYRMVHELPTFSSNPD 269
QY 121 YSELEHILAMTRDGEKTKARIRERVKSLSTPYPLSGENFRNN 164
DB 270 YSELEHILAMTRDGEKTKARIRERVKSLSTPYPLSGENFRNN 313

RESULT 4

AA682049
ID AAG82049 standard; Protein; 299 AA.

AC AAG82049;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:1192.

KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
KM vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimmerly WJ;

DR WPI: 2001-316495/33.

DR N-PSDB; AAH52899.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX useful for vaccinating against infections, e.g. endocarditis -

PS Claim 18; Page 342; 2188bp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide

CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 299 AA;

Query Match 91.9%; Score 782; DB 22; Length 299;
Best Local Similarity 89.6%; Pred. No. 1.1e-78;
Matches 147; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

QY 1 AADDTCTAATNGEQYKGLYLPGPFGTGSFIIIGALANOLKSKVNSTIYVLEPIRTLKG 60
DB 136 AADDTCTAATNGEQYKGLYLPGPFGTGSFIIIGALANOLKSKVNSTIYVLEPIRTLKG 195
QY 61 GFKDSFEKKLRRVREANITLMDIGAEVTPWVDEVIIGPLHYRMVHELPTFSSNPD 120
DB 196 GFKDSFEKKLRRVREANITLMDIGAEVTPWVDEVIIGPLHYRMVHELPTFSSNPD 255
QY 121 YSELEHILAMTRDGEKTKARIRERVKSLSTPYPLSGENFRNN 164
DB 256 YSELEHILAMTRDGEKTKARIRERVKSLSTPYPLSGENFRNN 299

RESULT 5

AA682337
ID AAG82337 standard; Protein; 299 AA.

AC AAG82337;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:11768.

KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
KM vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN WO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimmerly WJ;

DR WPI: 2001-316495/33.

DR N-PSDB; AAH53187.

PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX useful for vaccinating against infections, e.g. endocarditis -

PS Claim 18; Page 488; 2188bp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis.
CC (I) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used

in the exemplification of the present invention.
 N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4455 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472, no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 299 AA;

Query Match 91.9%; Score 782; DB 22; Length 299;
 Best Local Similarity 89.6%; Pred. No. 1,1e-78;
 Matches 147; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

QY 1 AADDICTAATNGEQVGLYLPGPGTGSFIIIGAIANQKSKKVRSTIYLPFIRTLKG 60
 DB 136 AADQICTAATNDKVKGLYIGPGTGSFIIIGAIANQKSKKSTIYLPFIRTLKG 195
 QY 61 GFGDGSFEKKLHVRANILMDLDIGAEVTPWVRDEVIGPLIHYRMVHELPTFFSSNPD 120
 DB 196 GFGDGSFEKKLQVRANILMDLDIGAEVTPWVRDEVIGPLIHYRMVHELPTFFSSNPN 255
 QY 121 YSELHHLMTDRDGEKTKAARIIEVKSLSTPYPLSGENPRNN 164
 DB 256 YSELHHLSTIRDTGTEKTKAARIIEKIKLSTPYPLTGKNPRNN 299

RESULT 6

ABP38947
 ID ABP38947 standard; Protein; 307 AA.

AC ABP38947;

DT 24-JUL-2002 (first entry)

DE Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3792.

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KM antibacterial; gene therapy.

XX Staphylococcus epidermidis.

XX US6380370-B1.

XX 30-APR-2002.

XX 13-AUG-1998; 98US-0134001.

XX 14-AUG-1997; 97US-055779P.

XX 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2002-381255/41.

XX N-PSDB; ABN91492.

XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections -

XX Disclosure; SEQ ID 3792; 267pp; English.

XX ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC N.B.: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.

XX Sequence 307 AA;

Query Match 91.9%; Score 782; DB 23; Length 307;
 Best Local Similarity 89.6%; Pred. No. 1,1e-78;
 Matches 147; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

QY 1 AADDICTAATNGEQVGLYLPGPGTGSFIIIGAIANQKSKKVRSTIYLPFIRTLKG 60
 DB 144 AADQICTAATNDKVKGLYIGPGTGSFIIIGAIANQKSKKSTIYLPFIRTLKG 203
 QY 61 GFGDGSFEKKLHVRANILMDLDIGAEVTPWVRDEVIGPLIHYRMVHELPTFFSSNPD 120
 DB 204 GFGDGSFEKKLQVRANILMDLDIGAEVTPWVRDEVIGPLIHYRMVHELPTFFSSNPN 263
 QY 121 YSELHHLMTDRDGEKTKAARIIEVKSLSTPYPLSGENPRNN 164
 DB 264 YSELHHLSTIRDTGTEKTKAARIIEKIKLSTPYPLTGKNPRNN 307

RESULT 7

ABB49355
 ID ABB49355 standard; Protein; 307 AA.

AC ABB49355;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #2059.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

XX WO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR01118.

XX 11-APR-2000; 2000FR-0004629.

XX (INSP) INST PASTEUR.

XX Buchrieser C, Frangoul L, Couve E, Ruenick C, Feihl H, Dehoux P,
 PI Dussurge O, Chetouani F, Nedjari H, Glaeser P, Kuntz F, Cosserat P,
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA,
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
 PI Chakraborty T, Domann E, Hain T, Berche P, Chardot A, Durant L,
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Madueno E, De Pablo B, Weiland J, Kaerst U, Entian K, Hauf J,
 PI Rose M, Voss H;

XX WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides -

XX Claim 6; SEQ ID NO 2060; 192pp; French.

XX The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABN03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication

CC in the specification). ABN15762 to ABN27252 encode the human ORFX

DR N-PSDB; ABN68460

CC in the specification). ABN15762 to ABN27252 encode the human ORFX

PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -

PS Claim 1; Page 3645; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71516 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.

XX Sequence 300 AA;

Query Match 41.2%; Score 351; DB 23; Length 300;

Best Local Similarity 43.1%; Pred. No. 1,3e-30;
 Matches 69; Conservative 36; Mismatches 53; Indels 2; Gaps 2;

QY 4 DICTAITNGEQVKGVLVYGGPGTGKSPFLGAINQLSKK-VNSTIYLDEPFRITLKGKF 62

DB 142 DPEQVPSAEQ-KGLVYGDVGIGKSYLLAMAHLESEKGVSTTLHPSPFAIDVGNAI 200

QY 63 KDSFEKKHRYVEANILMDIGAEVTPWVRDEVIGPLLHRYMHELPFSSNDYS 122

DB 201 SNCSVKEIDAVKVPFLIDIGAEQATSWRDEVQLVQLYRMLELPFTSNYSFA 260

QY 123 ELEHHLAMTRDGEKTKAARIIRVKSLSTPYFLSGENFR 162

DB 261 DLERKWAATIKGSDPTWQAKVMERVRYLAEPHLEGANRR 300

RESULT 10

ABU02187
 ID ABU02187 standard; Protein; 298 AA.

XX ABU02187;

XX 11-FEB-2003 (first entry)

XX 8. pneumoniae type 4 strain protein from coding region #1764.

XX Bacterial meningitis; pneumonia; sepsis; otitis media;

XX ear infection; antiinflammatory; antibacterial; immunostimulant;

XX audiotory; respiratory; gene therapy; vaccine.

OS Streptococcus pneumoniae type 4 strain.

XX WO200277021-A2.

PD 03-OCT-2002.

PF 27-MAR-2002; 2002WO-IB02163.

PR 27-MAR-2001; 2001GB-0007658.

PA (CHIR-) CHIRON SPA.

PI Masignani V, Tettelein H, Frazer C;

DR WPI; 2003-040579/03.

DR N-PSDB; ABX07476.

XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,

PT useful as medicaments for treating or preventing a disease or infection

PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media

PT or ear infection -

PS Claim 1; SEQ ID NO 3528; 56pp; English.

CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC AB556454. Also included are an antibody which binds one of the
 CC proteins, treating a patient by administering the protein, DNA or
 CC antibody (in a composition), a kit comprising first and second primers,
 CC which are the nucleic acid cited above or fragments between nucleotides
 CC 8-100 of a sequence not defined in the specification, for amplifying a
 CC target sequence contained within a Streptococcus nucleic acid sequence,
 CC where the first primer is substantially complementary to the target
 CC sequence and the second primer is substantially complementary to the
 CC complement of the target sequence, and where the parts of the primers
 CC having substantial complementarity define the termini of the target
 CC sequence to be amplified, assay comprising contacting a test compound
 CC with the protein, and determining whether the test compound binds to the
 CC protein and a Streptococcus pneumoniae bacterium, where one or more
 CC genes encoding the proteins has been rendered inactive. The proteins,
 CC nucleic acid molecules, antibody and compositions are useful as
 CC medicaments for treating or preventing a disease or infection due to
 CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
 CC sepsis, otitis media or ear infection. They are also useful in developing
 CC vaccines, diagnostics and antibiotics. The methods are useful for
 CC identifying immunodominant proteins. The present sequence is one of
 CC the 2469 proteins expressed by the identified coding regions from the
 CC genomic sequence.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 298 AA;

Query Match 40.7%; Score 346; DB 24; Length 298;

Best Local Similarity 44.3%; Pred. No. 4,8e-30;
 Matches 66; Conservative 35; Mismatches 46; Indels 2; Gaps 2;

QY 16 KGLVYGGPGTGKSPFLGAINQLSKK-VNSTIYLDEPFRITLKGFGSGFEKKHRY 74

DB 150 KGLVYGDVGIGKSYLLAMAHLESEKGVSTTLHPSPFAIDVGNAI SDGVNKTLYDBI 209

QY 75 REANILMDIGAEVTPWVRDEVIGPLLHRYMHELPFSSNDYSLEHHLAMTRDG 134

DB 210 KJSEVILLIDIGAEQATSWRDEVQLVQLYRMLELPFTSNFSEDLERKFAVYKRG 269

QY 135 EKKTKAARIIRVKSLSTPYFLSGENFR 162

DB 270 NDETWARVRYMERIRYLAETRLLEGVNR 298

RESULT 11

ABP27828
 ID ABP27828 standard; Protein; 300 AA.

XX ABP27828;

XX 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4832.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KW group A streptococcus; Streptococcus pyogenes; antibacterial;

KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

XX	Streptococcus agalactiae.
OS	
XX	WO200234771-A2.
PN	
XX	02-MAY-2002.
PD	
XX	29-OCT-2001; 2001MO-GB04789.
PF	
XX	27-OCT-2000; 2000GB-0026333.
KX	
PR	24-NOV-2000; 2000GB-0028727.
XX	07-MAR-2001; 2001GB-0005640.
PA	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
PI	Telford J, Maignani V, Margarit Ros YI, Grandt G, Fraser C;
PI	Rettelin H;
DR	WPI; 2002-352536/38.
DR	N-PSDB; ABN68459.
XX	
XX	New Streptococcus protein for the treatment or prevention of infection
PT	or disease caused by Streptococcus bacteria, such as meningitis, and
PT	for detecting a compound that binds to the protein -
PS	Claim 1; Page 3645; 4525pp; English.
CC	
XX	The invention relates to a protein (ABP25413-ABP30895) from group B
CC	streptococcus/BGS (Streptococcus agalactiae) or group A streptococcus/GAS
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC	the specification. The proteins have antibacterial and antiinflammatory
CC	activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
CC	antibodies that bind (1) are used in the manufacture of medicaments for
CC	the treatment or prevention of infection or disease caused by
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC	Nucleic acids encoding (1) are used to detect Streptococcus in a
CC	biological sample. (1) is used to determine whether a compound binds to
CC	(1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC	used as a vaccine or diagnostic composition. The disease caused by
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic
CC	acid encoding (1) may be used to recombinantly produce (1) and may be
CC	used in gene therapy. Antibodies to (1) are used for affinity
CC	chromatography, immunoassays, and distinguishing/identifying
CC	Streptococcus proteins.
XX	
SO	Sequence 300 AA:
	Query Match 39.1%; Score 332.5; DB 23; Length 300;
	Best Local Similarity 41.9%; Pred. No. 1.5e-28;
	Matches 62; Conservative 36; Mismatches 49; Indels 1; Gaps 1
OY	16 KGIIYGPRTGSGFLGAINQKSK VRSTIIYLEBFIRTLKGKGKGSPEKKAHRY 74
Db	153 KGIYLVDGDVGWGVSYYMAAMARLSRKKSFTLLHFSFAIDVNALISSTVDELDIAV 212
OY	75 RBNATIMLDDIGAEEVTMPWRDEVIGPLLAHYPMWELEPTFPSSNDYSLEHHLAMTRDG 134
Db	213 KSVPILLDDDIGAEQATSNVRDEILVOIHMMLESLPTFTSNYSFADLERKANIKGS 272
OY	135 EEXTKAARIIEKYSLSTPYFLSGENFR 162
Db	273 DETMQAKRWVERVRYLAIEFHLEGPRRR 300
RESULT 12	
ABBS4065	
ID	ABBS4065 standard; Protein; 293 AA.
AC	ABBS4065;
XX	
XX	16-MAY-2002 (first entry)
XT	

```

DE Lactococcus lactis protein dnal.
XX
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
XX Lactococcus lactis IL1403.
OS
FR2807446-A1.
XX
XX 12-OCT-2001.
PD
XX
XX 11-APR-2000; 2000FR-0004630.
PF
XX
PR 11-APR-2000; 2000FR-0004630.
XX
PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX
PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX
XX WPI; 2002-043416/06.
DR
XX
XX New nucleotide sequence useful in the identification or Lactococcus
PT Lactis and related species -
XX
PS Claim 6; SEQ ID No 767; 2504bp; French.
CC
XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (AB950521) and related proteins (AB953300-AB955621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO2001/77334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 293 AA;
Query Match 38.8%; Score 330; DB 23; Length 293;
Best Local Similarity 43.4%; Pred. No. 2,8e-28;
Matches 69; Conservative 30; Mismatches 50; Indels 10; Gaps 2;
QY 9 ITNGEQVKGVLVGPFGTGSFILGAIANOIKSKKVRSTIIYLPEFIRTLTKGFSDGSFE 68
DB 142 ITNPPKKGGIYLVGDPGVGKSFMLAAMANELAKKGIITLLHYPTFI-----SDLDFD 194
QY 69 KK--LHRVRANILIMDDIGAEEVTWVRDEVIGPLHYVMHLEPFFSSNFDSYLE 125
DB 195 NAKVWVEIKRSQVLVDDIGAEDNNMVAWRDIIQLVLIQHMOENLPTFTSNLMBELE 254
QY 126 HHLMTRDGESEKTKAARIERVKSLSPYPYLGSGENFRNN 164
DB 255 CHLAETKRADETWPAKRVMEKVYLAESMRLEGTRNRHD 293
RESULT 13
AAAG98965
ID AAG98965 standard; Protein; 248 AA.
XX
XX AAG98965;
XX
DT 26-SEP-2001 (first entry)
XX
DE E. coli growth and proliferation related protein sequence SEQ ID NO:435.
XX
XX Escherichia coli; growth; proliferation; microbial; antimicrobial;
KW bacterial infection; microorganism.
XX
OS Escherichia coli.
XX
FN WO200134810-A2.
XX
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:41 ; Search time 8.57221 Seconds
(without alignments)
809,474 Million cell updates/sec

Title: US-09-689-952-16

Sequence: 1 AADDICTAITNGEQVKGSLV.....ERKSLSTPYFLSGENFRNN 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	851	100.0	313	4	US-09-470-512A-2
2	849	99.8	313	4	US-09-470-512A-10
3	782	91.9	307	4	US-09-134-001C-3792
4	427.5	50.2	280	4	US-09-470-512A-14
5	420.5	49.4	278	4	US-09-470-512A-15
6	343.5	40.4	310	4	US-09-107-532A-5580
7	105.5	12.4	268	4	US-09-328-352-5963
8	102.5	12.0	304	4	US-09-134-001C-5153
9	90	10.6	389	2	US-08-820-170A-13
10	90	10.6	389	3	US-09-055-699-13
11	90	10.6	389	3	US-09-273-565-13
12	90	10.6	389	4	US-09-565-538-13
13	90	10.6	389	4	US-09-661-468-13
14	90	10.6	389	4	US-09-976-165-13
15	86.5	10.2	1349	2	US-08-612-734B-2
16	86	10.1	529	4	US-09-252-991A-18630
17	86	10.1	1334	2	US-08-596-545-2
18	86	10.1	1334	3	US-09-328-320-2
19	84	9.9	467	4	US-09-198-452A-443
20	83.5	9.8	456	3	US-09-268-364-21
21	83	9.8	308	4	US-09-149-476-607
22	83	9.8	399	4	US-09-149-476-474
23	83	9.8	806	3	US-08-999-774A-6
24	82	9.6	227	4	US-09-291-170A-9
25	81.5	9.6	227	4	US-09-724-884-9
26	81.5	9.6	266	4	US-09-218-363-8
27	81.5	9.6	267	4	US-09-218-363-10

28	81.5	9.6	290	4	US-09-218-363-4	Sequence 4, Appl1
29	81	9.5	220	4	US-09-790-179-4	Sequence 4, Appl1
30	80.5	9.5	264	4	US-09-328-352-7647	Sequence 7647, Ap
31	80.5	9.5	957	4	US-09-107-532A-4021	Sequence 4021, Ap
32	80	9.4	850	4	US-09-107-532A-4631	Sequence 4631, Ap
33	79.5	9.3	194	4	US-09-218-363-16	Sequence 16, Appl
34	79.5	9.3	134	4	US-09-218-363-17	Sequence 17, Appl
35	79.5	9.3	332	4	US-09-218-363-2	Sequence 2, Appl1
36	79.5	9.3	332	4	US-09-218-363-6	Sequence 6, Appl1
37	79.5	9.3	332	4	US-09-218-363-9	Sequence 9, Appl1
38	79.5	9.3	332	4	US-09-149-476-467	Sequence 467, App
39	79.5	9.3	859	4	US-09-199-637A-281	Sequence 281, App
40	79.5	9.3	859	4	US-09-252-991A-21413	Sequence 21413, A
41	78.5	9.2	320	4	US-09-252-991A-32024	Sequence 32024, A
42	78	9.2	215	4	US-09-291-170A-6	Sequence 6, Appl1
43	78	9.2	215	4	US-09-724-884-6	Sequence 6, Appl1
44	78	9.2	226	3	US-09-268-364-2	Sequence 2, Appl1
45	78	9.2	405	2	US-08-222-719-2	Sequence 2, Appl1

ALIGNMENTS

```

RESULT 1
US-09-470-512A-2
; Sequence 2, Application US/09470512A
; Patent No. 6376652
; GENERAL INFORMATION:
; APPLICANT: PhageTech, Inc.
; TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
; TITLE OF INVENTION: gene and its encoded protein
; FILE REFERENCE: 21715/1000
; CURRENT APPLICATION NUMBER: US/09/470,512A
; CURRENT FILING DATE: 1999-12-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-470-512A-2

Query Match      100.0%; Score 851; DB 4; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.1e-96;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AADDICTAITNGEQVKGSLVGPFGTKSFLIGALANOLSKKVSSTIYLPEFIRTLKG 60
DB      150 AADDICTAITNGEQVKGSLVGPFGTKSFLIGALANOLSKKVSSTIYLPEFIRTLKG 209
QY      61 GPKDGFPEKCLHVRVREANILMDITGAEVTPWVDEVIQPLHTRMVELPTFFSSNFD 120
DB      210 GPKDGFPEKCLHVRVREANILMDITGAEVTPWVDEVIQPLHTRMVELPTFFSSNFD 269
QY      121 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFLSGENFRNN 164
DB      270 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFLSGENFRNN 313

RESULT 2
US-09-470-512A-10
; Sequence 10, Application US/09470512A
; Patent No. 6376652
; GENERAL INFORMATION:
; APPLICANT: PhageTech, Inc.
; TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
; TITLE OF INVENTION: gene and its encoded protein
; FILE REFERENCE: 21715/1000
; CURRENT APPLICATION NUMBER: US/09/470,512A
; CURRENT FILING DATE: 1999-12-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10

```

LENGTH: 313
TYPE: PR
ORGANISM: Staphylococcus aureus
US-09-470-512A-10

Query Match 99.8%; Score 849; DB 4; Length 313;
Best Local Similarity 99.4%; Pred. No. 3.8e-98;
Matches 163; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AADICTAINGEYVGLYGGPGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 60
DB 150 AADICTAINGEYVGLYGGPGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 209
QY 61 GFDGSEFKLHRYREANIIMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFD 120
DB 210 GFDGSEFKLHRYREANIIMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFD 269
QY 121 YSELEHNLAMTRDGEETKARIIERYKSLSTPYFLSGENFRNN 164
DB 270 YSELEHNLAMTRDGEETKARIIERYKSLSTPYFLSGENFRNN 313

RESULT 3
US-09-134-001C-3792
Sequence 3792, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GFC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3792
LENGTH: 307
TYPE: PR
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3792

Query Match 91.9%; Score 782; DB 4; Length 307;
Best Local Similarity 89.6%; Pred. No. 9.5e-90;
Matches 147; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

QY 1 AADICTAINGEYVGLYGGPGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 60
DB 144 AADICTAINGEYVGLYGGPGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 203
QY 61 GFDGSEFKLHRYREANIIMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFD 120
DB 204 GFDGSEFKLHRYREANIIMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFD 263
QY 121 YSELEHNLAMTRDGEETKARIIERYKSLSTPYFLSGENFRNN 164
DB 264 YSELEHNLAMTRDGEETKARIIERYKSLSTPYFLSGENFRNN 307

RESULT 4
US-09-470-512A-14
Sequence 14, Application US/09470512A
Patent No. 6376652
GENERAL INFORMATION:
APPLICANT: PhageTech, Inc.
TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
FILE REFERENCE: 21715/1000
CURRENT APPLICATION NUMBER: US/09/470,512A
CURRENT FILING DATE: 1999-12-12
NUMBER OF SEQ ID NOS: 16

SOFTWARE: Patentin version 3.0
SEQ ID NO 14
LENGTH: 280
TYPE: PR
ORGANISM: Bacillus subtilis
US-09-470-512A-14

Query Match 50.2%; Score 427.5; DB 4; Length 280;
Best Local Similarity 55.7%; Pred. No. 2.7e-45;
Matches 83; Conservative 26; Mismatches 39; Indels 1; Gaps 1;

QY 16 KGLYVGPPTGSKSFLGAIANOLSKKVRSTIYLPFIRTLKGFDGSEFKLHRYR 75
DB 132 KGLYVGPPTGSKSFLGAIANOLSKKVRSTIYLPFIRTLKGFDGSEFKLHRYR 191
QY 76 EANIIMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFDYSELEHNLAMTRDGE 135
DB 192 TPEVLMDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFDYSELEHNLAMTRDGE 251
QY 136 -EKTARIIERYKSLSTPYFLSGENFRN 163
DB 252 KEEVKARLIERLILYLAIRLDGENRRH 280

RESULT 5
US-09-470-512A-15
Sequence 15, Application US/09470512A
Patent No. 6376652
GENERAL INFORMATION:
APPLICANT: PhageTech, Inc.
TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
FILE REFERENCE: 21715/1000
CURRENT APPLICATION NUMBER: US/09/470,512A
CURRENT FILING DATE: 1999-12-12
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patentin version 3.0
SEQ ID NO 15
LENGTH: 278
TYPE: PR
ORGANISM: Bacillus halodurans
US-09-470-512A-15

Query Match 49.4%; Score 420.5; DB 4; Length 278;
Best Local Similarity 50.0%; Pred. No. 2e-44;
Matches 77; Conservative 33; Mismatches 43; Indels 1; Gaps 1;

QY 12 GEQVGLYVGPPTGSKSFLGAIANOLSKKVRSTIYLPFIRTLKGFDGSEFKL 71
DB 125 GEQVGLYVGPPTGSKSFLGAIANOLSKKVRSTIYLPFIRTLKGFDGSEFKL 184
QY 72 HRYREANIIMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFDYSELEHNLAMTR 131
DB 185 DFKNAQVILFDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFDYSELEHNLAMTR 244
QY 132 -RDGEETKARIIERYKSLSTPYFLSGENFRNN 164
DB 245 DKSGETLAKRVERIRIRHTVSVWVGQVRREH 278

RESULT 6
US-09-107-532A-5580
Sequence 5580, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham

STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5580:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...310
SEQUENCE DESCRIPTION: SEQ ID NO: 5580:
US-09-107-532A-5580
Query Match 40.4%; Score 343.5; DB 4; Length 310;
Best Local Similarity 46.3%; Pred. No. 1.1e-34;
Matches 69; Conservative 28; Mismatches 51; Indels 1; Gaps 1;
QY 16 KELYVGPFGTGSFLGAIANOLSKKVRSTIYLPFRITLKGSGFKDSEKLLHYR 75
DB 161 KELYVSGIGIKYTLGATARELAAGTTTLVHPTFAVEKKAIGKQYAEKLDAYK 220
QY 76 EANIIMLDIGAEVYTPWVRDEVIGPLHYRWVHELPTFFSSNFDYSELEHHLAMTRDG- 134
DB 221 RPIIMIDDIGADAMSWIRDEFVGLQYRMQEQLPFTSSNFTMDLEQHLSTVQRGD 280
QY 135 EKTQAARIIEVKSLSTPYPLSGENFRN 163
DB 281 EEPKAKRIMERIRIRYLKEIEMTGRNRN 309
RESULT 7
US-09-328-352-5963
Sequence 5963, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 5963
LENGTH: 268
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5963

Query Match 12.4%; Score 105.5; DB 4; Length 268;
Best Local Similarity 27.0%; Pred. No. 6.9e-05;
Matches 24; Conservative 19; Mismatches 41; Indels 5; Gaps 2;
QY 2 ADDICTAI--TNGEQVGLVYGPFGTGSFLGAIANOLSKKVRSTIYLP--PEFIR 56
DB 143 AAEFCRKYLTQLGASQHNPLFLYGPGLGKTHLMQAVGNALLQAKPNARVMYMTSQSFVQ 202
QY 57 TLKGGFKDGSPEKLLHRRREANIIMLDI 85
DB 203 DFVSSLQKQKEEFKRCNSLDLLVDDI 231
RESULT 8
US-09-134-001C-5153
Sequence 5153, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5153
LENGTH: 304
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5153
Query Match 12.0%; Score 102.5; DB 4; Length 304;
Best Local Similarity 21.4%; Pred. No. 0.0002;
Matches 33; Conservative 35; Mismatches 59; Indels 27; Gaps 4;
QY 2 ADDICTAITNGEQVGLVYGPFGTGSFLGAIANOLSKKVRSTIYLP-----EPI 55
DB 135 AASLVAEAPAPAAVYPLFTYGVGLGKTHLMALIGHVLSNPNKAVIYTSSEKFTNEFI 194
QY 56 RTLKGGFKDGSPEKLLHRRREANIIMLDIGAEVTPWVRDEVIGPLHYRWVHELPTFF 115
DB 195 KSI---BDNETAEFRERYKIDVLLIDIDIOFTQKQTOEE-----FF 234
QY 116 SSNFDYSELEHHLAMTRDGEKTKAARIIEVKYS 149
DB 235 HTFNEHONKQIYISSDRPK-EIAKLEDRLRS 267
RESULT 9
US-08-820-170A-13
Sequence 13, Application US/08820170A
Patent No. 5631058
GENERAL INFORMATION:
APPLICANT: Tautomu, FUMIYARA
APPLICANT: Takeshi, WATANABE
APPLICANT: Masato, HORIE
APPLICANT: Toyomasa, KATAGIRI
TITLE OF INVENTION: HUMAN GENE
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sughrue, Mion, Zimm, Macpeak & Seas
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: United States
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/820,170A
 FILING DATE:
 CLASSIFICATION: 536
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 293-7060
 TELEFAX: (202) 293-7860
 TELEEX: 6491103
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 389 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-820-170A-13

Query Match 10.6%; Score 90; DB 2; Length 389;
 Best Local Similarity 35.1%; Pred. No. 0.011;
 Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLIYGPFGTGSFLGAIANOLKS--KKVRSTII--YLPEFIRTLKGFPGDSFEKK 70
 DB 168 KGLIYGPFGTGTKLARAVASQLDCNPLKVVSSIVDKYIGESARLIREFM-----N 220

QY 71 LHRVREANTIMLDDIGA 87
 DB 221 YARDHQPCLIFMDEIDA 237

RESULT 10
 US-09-055-699-13
 Sequence 13, Application US/09055699
 Patent No. 6005088
 GENERAL INFORMATION:
 APPLICANT: Teutomu, FUJIWARA
 APPLICANT: Takeshi, WATANABE
 APPLICANT: Masato, HORIE
 APPLICANT: Toyomasa, KATAGIRI
 TITLE OF INVENTION: HUMAN GENE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
 STREET: 2100 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States
 ZIP: 20037-3202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/055,699
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/820,170
 FILING DATE:
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 293-7060
 TELEFAX: (202) 293-7860
 TELEEX: 6491103
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 389 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-055-699-13

Query Match 10.6%; Score 90; DB 3; Length 389;
 Best Local Similarity 35.1%; Pred. No. 0.011;
 Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLIYGPFGTGSFLGAIANOLKS--KKVRSTII--YLPEFIRTLKGFPGDSFEKK 70
 DB 168 KGLIYGPFGTGTKLARAVASQLDCNPLKVVSSIVDKYIGESARLIREFM-----N 220

QY 71 LHRVREANTIMLDDIGA 87
 DB 221 YARDHQPCLIFMDEIDA 237

RESULT 11
 US-09-273-565-13
 Sequence 13, Application US/09273565A
 Patent No. 616190
 GENERAL INFORMATION:
 APPLICANT: FUJIWARA, TSUTOMU
 APPLICANT: WATANABE, TAKESHI
 APPLICANT: HORIE, MASATO
 TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
 FILE REFERENCE: Q-53599
 CURRENT APPLICATION NUMBER: US/09/273,565A
 CURRENT FILING DATE: 1999-03-22
 EARLIER APPLICATION NUMBER: 09/055,699
 EARLIER FILING DATE: 1998-04-07
 EARLIER APPLICATION NUMBER: 08/820,170
 EARLIER FILING DATE: 1997-03-19
 EARLIER APPLICATION NUMBER: JP 63410/1996
 EARLIER FILING DATE: 1996-03-19
 EARLIER APPLICATION NUMBER: JP 69163/1997
 EARLIER FILING DATE: 1997-03-05
 NUMBER OF SEQ ID NOS: 95
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 13
 LENGTH: 389
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-273-565-13

Query Match 10.6%; Score 90; DB 3; Length 389;
 Best Local Similarity 35.1%; Pred. No. 0.011;
 Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLIYGPFGTGSFLGAIANOLKS--KKVRSTII--YLPEFIRTLKGFPGDSFEKK 70
 DB 168 KGLIYGPFGTGTKLARAVASQLDCNPLKVVSSIVDKYIGESARLIREFM-----N 220

QY 71 LHRVREANTIMLDDIGA 87
 DB 221 YARDHQPCLIFMDEIDA 237

RESULT 12
 US-09-565-538-13
 Sequence 13, Application US/09565538
 Patent No. 633404
 GENERAL INFORMATION:
 APPLICANT: FUJIWARA, TSUTOMU
 APPLICANT: WATANABE, TAKESHI
 APPLICANT: HORIE, MASATO
 TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
 FILE REFERENCE: Q-53599
 CURRENT APPLICATION NUMBER: US/09/565,538
 CURRENT FILING DATE: 2000-05-05
 PRIOR APPLICATION NUMBER: 09/273,565
 PRIOR FILING DATE: 1999-03-22
 PRIOR APPLICATION NUMBER: 09/055,699
 PRIOR FILING DATE: 1998-04-07
 PRIOR APPLICATION NUMBER: 08/820,170

PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 389
TYPE: PRT
ORGANISM: Homo sapiens
US-09-565-538-13

Query Match
Best Local Similarity 10.6%; Score 90; DB 4; Length 389;
Best Local Similarity 35.1%; Pred. No. 0.011;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLVYGPFGTGSFLIGAIANOLKS---KVRSTII--YLPEFIRTLKGFDSFEKK 70
DB 168 KGLVYGPFGTGTTLARAVASQDCNFKVSSIVDKYIGSARLIREMF-----N 220

QY 71 LHRVREANILMDIGA 87
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 13
US-09-661-468-13
Sequence 13, Application US/09661468
Patent No. 6376189
GENERAL INFORMATION:
APPLICANT: FUJIMURA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/661,468
CURRENT FILING DATE: 2000-09-13
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 389
TYPE: PRT
ORGANISM: Homo sapiens
US-09-661-468-13

Query Match
Best Local Similarity 10.6%; Score 90; DB 4; Length 389;
Best Local Similarity 35.1%; Pred. No. 0.011;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLVYGPFGTGSFLIGAIANOLKS---KVRSTII--YLPEFIRTLKGFDSFEKK 70
DB 168 KGLVYGPFGTGTTLARAVASQDCNFKVSSIVDKYIGSARLIREMF-----N 220

QY 71 LHRVREANILMDIGA 87
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 14
US-09-976-165-13
Sequence 13, Application US/09976165
Patent No. 6562947
GENERAL INFORMATION:

APPLICANT: FUJIMURA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/976,165
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/565,538
PRIOR FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 13
LENGTH: 389
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-165-13

Query Match
Best Local Similarity 10.6%; Score 90; DB 4; Length 389;
Best Local Similarity 35.1%; Pred. No. 0.011;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLVYGPFGTGSFLIGAIANOLKS---KVRSTII--YLPEFIRTLKGFDSFEKK 70
DB 168 KGLVYGPFGTGTTLARAVASQDCNFKVSSIVDKYIGSARLIREMF-----N 220

QY 71 LHRVREANILMDIGA 87
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 15
US-08-612-734B-2
Sequence 2, Application US/08612734B
Patent No. 5914246
GENERAL INFORMATION:
APPLICANT: Peery, Robert B.
APPLICANT: Skatrud, Paul L.
APPLICANT: Tobin, Matthew B.
TITLE OF INVENTION: Multiple Drug Resistance Gene, of
TITLE OF INVENTION: Aspergillus Fumigatus
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center, DC1501
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,734B
FILING DATE: 08-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Craig, Anne I.
REGISTRATION NUMBER: 32,976
REFERENCE/DOCKET NUMBER: X-9681
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-354-9570
TELEFAX: 617-354-4043

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-612-734B-2

Query Match 10.2%; Score 86.5; DB 2; Length 1349;
Best Local Similarity 20.9%; Pred. No. 0.2;
Matches 32; Conservative 27; Mismatches 49; Indels 45; Gaps 7;

Qy	3	DDICTAITNGEQVKGVLVYGFPGTGKSPILGAIANQLSKKVRSTIYLPPIRTLKGF	62
Db	456	EDVSLSMFAG---KTTALVGPSPSGSKSTVGLVER-----FYLPGGQVLLDG-	500
Qy	63	KDGSFEKKLHRVREANILMLDDIGAEEVTPWVRDEVIGPILHYRMVHELPTFPSSNPDYS	122
Db	501	-----HDIQTLNL-----RWLRQOI-----SLVSOEPVLPSTTI-FR	531
Qy	123	ELBHLAMTR-DGEKTKAARIIRVKSLSPTY	154
Db	532	NIEHGLIGTKFEHESKDKIRELVENAMAMNAH	564

Search completed: December 15, 2003, 15:17:39
Job time : 8.57221 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:13:51 / Search time 16.0165 Seconds
(without alignments)
1904.368 Million cell updates/sec

Title: US-09-689-952-16
Perfect score: 851
Sequence: 1 AADICTAINTGEQVKGILY.....ERVKSLSTPYPLSGENFRNN 164

Scoring table: BLOSUM62
Gapop 10.0, Gapept 0.5

Searched: 684280 seqs, 185983659 residues
Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database: Published Applications AA.*
- 1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubppaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104.5	12.3	248	US-10-287-274-435	Sequence 435, App
2	100	11.8	653	US-10-156-761-11850	Sequence 11850, A
3	98	11.5	467	US-09-815-242-10395	Sequence 10395, A
4	95.5	11.2	442	US-09-815-242-13552	Sequence 13552, A
5	92.5	10.9	574	US-09-764-868-774	Sequence 774, App
6	91.5	10.6	827	US-10-032-585-7621	Sequence 7621, App
7	90.5	10.6	454	US-09-815-242-11167	Sequence 11167, A
8	90	10.6	389	US-09-976-165-113	Sequence 11, App
9	90	10.6	389	US-10-205-219-165	Sequence 165, App
10	90	10.6	389	US-10-342-276-113	Sequence 13, App1
11	90	10.6	393	US-10-128-714-3421	Sequence 3421, App
12	90	10.6	451	US-10-128-714-8421	Sequence 8421, App
13	88.5	10.4	796	US-10-032-585-7201	Sequence 7201, App
14	88	10.3	444	US-10-032-585-7385	Sequence 7385, App
15	88	10.3	453	US-09-815-242-5668	Sequence 5668, App

16	88	10.3	453	9	US-09-815-242-12701	Sequence 12701, A
17	87	10.2	411	12	US-10-032-585-7284	Sequence 7284, App
18	86.5	10.2	606	15	US-10-128-714-3064	Sequence 3064, App
19	86.5	10.2	847	15	US-10-128-714-8064	Sequence 8064, App
20	86	10.1	439	15	US-10-128-714-3522	Sequence 3522, App
21	86	10.1	439	15	US-10-128-714-8522	Sequence 8522, App
22	86	10.1	514	9	US-09-815-242-11772	Sequence 11772, A
23	86	10.1	1334	10	US-09-758-828-2	Sequence 2, App1
24	85.5	10.0	444	15	US-10-153-668-308	Sequence 308, App
25	85.5	10.0	465	9	US-09-925-302-771	Sequence 771, App
26	85.5	10.0	482	10	US-09-764-868-1149	Sequence 1149, App
27	85.5	10.0	489	15	US-09-764-868-776	Sequence 776, App
28	85	10.0	421	15	US-10-128-714-8485	Sequence 8485, App
29	85	10.0	465	15	US-10-128-714-3485	Sequence 3485, App
30	84.5	9.9	442	11	US-09-828-062-7	Sequence 7, App1
31	84	9.9	401	12	US-10-032-585-7325	Sequence 7325, App
32	84	9.9	561	15	US-10-136-761-13374	Sequence 13374, A
33	84	9.9	826	12	US-10-032-585-7227	Sequence 7227, App
34	83.5	9.8	298	12	US-10-380-558-17	Sequence 17, App1
35	83.5	9.8	456	9	US-09-756-998-21	Sequence 21, App1
36	83	9.8	308	12	US-09-809-391-607	Sequence 607, App
37	83	9.8	308	12	US-09-882-171-607	Sequence 607, App
38	83	9.8	389	15	US-10-128-714-3369	Sequence 3369, App
39	83	9.8	389	15	US-10-128-714-8369	Sequence 8369, App
40	83	9.8	399	11	US-09-809-391-474	Sequence 474, App
41	83	9.8	329	12	US-09-882-171-474	Sequence 474, App
42	83	9.8	728	10	US-09-712-363-167	Sequence 167, App
43	83	9.8	806	9	US-09-922-217-1117	Sequence 1117, App
44	83	9.8	806	9	US-09-922-217-1118	Sequence 1118, App
45	83	9.8	806	12	US-10-341-434-99	Sequence 99, App1

ALIGNMENTS

RESULT 1
US-10-287-274-435
Sequence 435, Application US/10287274
Publication No. US20030181408A1
GENERAL INFORMATION:
APPLICANT: Foreyth, R. Allyn
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERETO
FILE REFERENCE: ELITRA, 008DVI
CURRENT APPLICATION NUMBER: US/10/287,274
CURRENT FILING DATE: 2002-10-31
PRIOR APPLICATION NUMBER: US 60/164415
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: US 09/711164
NUMBER OF SEQ ID NOS: 469
SOFTWARE: FaSTSeq for Windows Version 4.0
SEQ ID NO 435
LENGTH: 248
TYPE: PRT
ORGANISM: Escherichia coli
US-10-287-274-435
Query Match 12.3%; Score 104.5; DB 12; Length 248;
Best Local Similarity 27.2%; Pred. No. 0.00078;
Matches 40; Conservative 28; Mismatches 60; Indels 19; Gaps 6;
QY 2 AADICTAINTGEQVKGILYYPFGGKSTILAIANOLKSKVRSITITLPERIRLKG 61
DB 94 ADELMTGCTN-----FAFGSKRGTKGNHIAIGRLKDCQTVVVVADVMALHRS 147
QY 62 FPDG-SFEKKHRYEANTLMDDIGAEVTPWVDEVTGPIIHYRMVHELPTFFSSND 120
DB 148 YPDGSGEFTLELCEVDLVADEIGQRET---KNEQV-VLH-QIVRRIRASMSVGM 201
QY 121 VSELEHILMTDGEKTKVARIIRV 147

Db 202 LTNLYEAMKTLIGE-----RIMDRM 222

RESULT 2
US-10-156-761-11850

Sequence 11850, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

PRIOR FILING DATE: 2002-05-29

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 11850

LENGTH: 653

TYPE: PR

ORGANISM: Streptomyces avermitilis

US-10-156-761-11850

Query Match 11.8%; Score 100; DB 15; Length 653;

Best Local Similarity 39.2%; Pred. No. 0.01;

Matches 29; Conservative 15; Mismatches 20; Indels 10; Gaps 4;

QY 18 LTVYGPFGKSFILGAINOLKSKKVVSTIYLP--PEPRTLKGGFKDGSFEKKLHVR 71

DB 350 LFTYGGSGLGKTHLNAIGHVARSLYPTVRVYSSEEFNFINSGIRDG-KDSDPRK-- 406

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

QY 72 HVRREANITLDDI 85

DB 407 -RREMDILLVDI 419

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10395

LENGTH: 467

TYPE: PR

ORGANISM: Escherichia coli

US-09-815-242-13552

Query Match 11.5%; Score 98; DB 9; Length 467;

Best Local Similarity 27.1%; Pred. No. 0.011;

Matches 19; Conservative 21; Mismatches 28; Indels 2; Gaps 1;

QY 18 LTVYGPFGKSFILGAINOLKSKKVVSTIYLP--PEPRTLKGGFKDGSFEKKLHVR 75

DB 168 LFTYGGSGLGKTHLNAIGHVARSLYPTVRVYSSEEFNFINSGIRDG-KDSDPRK-- 227

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

QY 76 EANITLDDI 85

DB 228 SYDALIIDI 237

Db 200 SLDLLIDDIOS-----LSGKVVATQEEFFNTFNALHDKQKQIVLTSDRS 244

Qy 136 BK 137
Db 245 PK 246

RESULT 5
US-09-764-868-774

; Sequence 774, Application US/09764868
; Patent No. US20020168711A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT232

; CURRENT APPLICATION NUMBER: US/09/764, 868

; PRIOR FILING DATE: 2001-01-17

; PRIOR APPLICATION DATA REMOVED - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 774

; LENGTH: 574

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-868-774

Query Match 10.9%; Score 92.5; DB 10; Length 574;
Best Local Similarity 30.3%; Pred. No. 0.064;
Matches 30; Conservative 17; Mismatches 33; Indels 19; Gaps 4;

Qy 16 KGLVYGPFGTGSFLIGAIANOLKSKVRSTIYLPFRITLKGFGKDSFEKLRHV-74

Db 127 RGLVLYGPFGTGSFLIGAIANOLKSKVRSTIYLPFRITLKGFGKDSFEKLRHV-74

Qy 75 -----REANILMLDDIGA-----EYVTPWVRDEVIGPL 103

Db 179 AEATLHPSIIFIDELDALCPKREGAQNEVEKRVASLT 217

RESULT 6
US-10-032-585-7621

; Sequence 7621, Application US/10032585

; Publication No. US20030180953A1

; GENERAL INFORMATION:

; APPLICANT: Terry, Roemer D.

; APPLICANT: Bo, Jiang

; APPLICANT: Charles, Boone

; APPLICANT: Howard, Bussey

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery

; FILE REFERENCE: 10182-005-999

; CURRENT APPLICATION NUMBER: US/10/032, 585

; PRIOR FILING DATE: 2001-12-20

; NUMBER OF SEQ ID NOS: 8000

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 7621

; LENGTH: 827

; TYPE: PRT

; ORGANISM: Candida albicans

US-10-032-585-7621

Query Match 10.8%; Score 91.5; DB 12; Length 827;
Best Local Similarity 31.8%; Pred. No. 0.14; Indels 17; Gaps 6;
Matches 35; Conservative 17; Mismatches 41; Indels 17; Gaps 6;

Qy 10 TNGEQVKGVLVYGPFGTGSFLIGAIANOLKSKVRSTIYLPFRITLKGFGKDSFEK 69

Db 222 TNGEPRGVLYVPGPGKTTINALAGELKVFIN---ISASVSGM-----SGESRK 273

Qy 70 KLRHV-REAN-----ILMLDDIGAEEVTPWVRDEVIGPLHVRMVEHLP 113

Db 274 KLRHFEKAKQIAPCLIFMDEIDA--ITP-KRDGAQRERKRIVAQLLT 320

RESULT 7
US-09-815-242-11167

; Sequence 11167, Application US/09815242
; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haseibeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206, 848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207, 727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242, 578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253, 625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257, 931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269, 308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: RatsSeq for Windows Version 4.0

; SEQ ID NO 11167

; LENGTH: 454

; TYPE: PRT

; ORGANISM: Haemophilus influenzae

US-09-815-242-11167

Query Match 10.6%; Score 90.5; DB 9; Length 454;
Best Local Similarity 26.8%; Pred. No. 0.08;
Matches 37; Conservative 25; Mismatches 47; Indels 29; Gaps 7;

Qy 19 YVYGPFGTGSFLIGAIANOLKSKVRSTIYLPFRITLKGFGKDSFEK-KLHVRV 74

Db 157 FLYGTGKGLTHLHAGIADKPNKRVLYIHANNFQHVAKVRDKMDQFKKPYRS 216

Qy 75 REANILMLDDI---GAEEVTPWVRDEVIGPLHVRMVEHLPFSSNPFYSELEHILAMT 131

Db 217 LDA--LLVDDIQFPAKEKTKQ-----EEFFHIFNSLP---ETGRQIILT 255

Qy 132 RDGEETKARILIERVKS 149

Db 256 SDRVPK-EIEKIEERLKS 272

RESULT 8
US-09-976-165-13

; Sequence 13, Application US/09976165
; Patent No. US20020107383A1

; GENERAL INFORMATION:

; APPLICANT: FUJIMURA, TSUTOMU

; APPLICANT: WATANABE, TAKESHI

; APPLICANT: HORIE, MASATO

; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN

; FILE REFERENCE: O-63599

; CURRENT APPLICATION NUMBER: US/09/976, 165

; PRIOR FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: 09/565, 538

; PRIOR FILING DATE: 2000-05-05

; PRIOR APPLICATION NUMBER: 09/055, 699

;; PRIOR FILING DATE: 1998-04-07
;; PRIOR APPLICATION NUMBER: 08/820,170
;; PRIOR FILING DATE: 1997-03-19
;; PRIOR APPLICATION NUMBER: JP 63410/1996
;; PRIOR FILING DATE: 1996-03-19
;; PRIOR APPLICATION NUMBER: JP 69163/1997
;; PRIOR FILING DATE: 1997-03-05
;; NUMBER OF SEQ ID NOS: 95
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 13
;; LENGTH: 389
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-976-165-13

Query Match 10.6%; Score 90; DB 10; Length 389;
Best Local Similarity 35.1%; Pred. No. 0.073;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLYLGPGFTGKSFILGAIANOLKS---KVRSTII--YLPEFIRTLKGGFQDGSFEKK 70
DB 168 KGCILYGPFGTGTLLARAVASQDCNFKLVVSSIVDKYIGESARLIREMF-----N 220

QY 71 LHRVREANILMLDDIGA 87
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 9
US-10-205-219-165
;; Sequence 165, Application US/10205219
;; Publication No. US20030138803A1
;; GENERAL INFORMATION:
;; APPLICANT: Warner-Lambert Company
;; APPLICANT: Lee, Kevin
;; APPLICANT: Dixon, Alister
;; APPLICANT: Brookbank, Robert
;; APPLICANT: Plimock, Robert
;; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
;; FILE REFERENCE: WU-A-018200
;; CURRENT APPLICATION NUMBER: US/10/205,219
;; CURRENT FILING DATE: 2002-07-24
;; PRIOR APPLICATION NUMBER: GB 0118354.0
;; PRIOR FILING DATE: 2001-07-27
;; NUMBER OF SEQ ID NOS: 197
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 165
;; LENGTH: 389
;; TYPE: PRT
;; ORGANISM: *Spermophilus tridecemlineatus*
;; FEATURE:
;; OTHER INFORMATION: ATPase domain protein 44
US-10-205-219-165

Query Match 10.6%; Score 90; DB 12; Length 389;
Best Local Similarity 35.1%; Pred. No. 0.073;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLYLGPGFTGKSFILGAIANOLKS---KVRSTII--YLPEFIRTLKGGFQDGSFEKK 70
DB 168 KGCILYGPFGTGTLLARAVASQDCNFKLVVSSIVDKYIGESARLIREMF-----N 220

QY 71 LHRVREANILMLDDIGA 87
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 10
US-10-342-276-13
;; Sequence 13, Application US/10342276
;; Publication No. US20030143688A1
;; GENERAL INFORMATION:
;; APPLICANT: FUJIMURA, TSUTOMU

;; APPLICANT: WATANABE, TAKESHI
;; APPLICANT: HORIE, MASAO
;; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIOUITIN-CONUGATING ENZYME
;; FILE REFERENCE: Q-53599
;; CURRENT APPLICATION NUMBER: US/10/342,276
;; CURRENT FILING DATE: 2003-01-15
;; PRIOR APPLICATION NUMBER: US/09/976,165
;; PRIOR FILING DATE: 2001-10-15
;; PRIOR APPLICATION NUMBER: 09/565,538
;; PRIOR FILING DATE: 2000-05-05
;; PRIOR APPLICATION NUMBER: 09/055,699
;; PRIOR FILING DATE: 1998-04-07
;; PRIOR APPLICATION NUMBER: 08/820,170
;; PRIOR FILING DATE: 1997-03-19
;; PRIOR APPLICATION NUMBER: JP 63410/1996
;; PRIOR FILING DATE: 1996-03-19
;; PRIOR APPLICATION NUMBER: JP 69163/1997
;; PRIOR FILING DATE: 1997-03-05
;; NUMBER OF SEQ ID NOS: 95
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 13
;; LENGTH: 389
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-342-276-13

Query Match 10.6%; Score 90; DB 12; Length 389;
Best Local Similarity 35.1%; Pred. No. 0.073;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 16 KGLYLGPGFTGKSFILGAIANOLKS---KVRSTII--YLPEFIRTLKGGFQDGSFEKK 70
DB 168 KGCILYGPFGTGTLLARAVASQDCNFKLVVSSIVDKYIGESARLIREMF-----N 220

QY 71 LHRVREANILMLDDIGA 87
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 11
US-10-128-714-3421
;; Sequence 3421, Application US/10128714
;; Publication No. US20030119013A1
;; GENERAL INFORMATION:
;; APPLICANT: Jiang, Bo
;; APPLICANT: Hu, Weng
;; APPLICANT: Tishkoff, Daniel
;; APPLICANT: Zamudio, Carlos
;; APPLICANT: Broshkin, Alexey M
;; APPLICANT: Lemieux, Sebastien M
;; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
;; TITLE OF INVENTION: Methods of Use
;; FILE REFERENCE: 10162-018-999
;; CURRENT APPLICATION NUMBER: US/10/128,714
;; CURRENT FILING DATE: 2002-04-23
;; PRIOR APPLICATION NUMBER: US 60/285,697
;; PRIOR FILING DATE: 2001-04-23
;; PRIOR APPLICATION NUMBER: US 60/287,066
;; PRIOR FILING DATE: 2001-04-27
;; PRIOR APPLICATION NUMBER: US 60/295,890
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: US 60/303,899
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: US 60/316,362
;; PRIOR FILING DATE: 2001-08-31
;; NUMBER OF SEQ ID NOS: 8603
;; SOFTWARE: Patentin version 3.1
;; SEQ ID NO 3421
;; LENGTH: 393
;; TYPE: PRT
;; ORGANISM: *Aspergillus fumigatus*
US-10-128-714-3421

RESULT 13
US-10-032-585-7201
Sequence 7201, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, JIANG
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585

RESULT 15
US-09-815-242-5668
Sequence 5668, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard

```

; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5668
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-5668

Query Match      10.3%; Score 88; DB 9; Length 453;
Best Local Similarity 28.4%; Pred. No. 0.16;
Matches 21; Conservative 17; Mismatches 26; Indels 10; Gaps 2;

QY      18 LNYGPGTGKSPFLGAIANQLSKKVRSTIYLP-----EPIRTLKGGFKDGSPEKKL 71
      152 LFTYGGVGLGKTHLMAIGHVLDNNPDAKVITYTSSEKFTNEFIKSI---RDNEGGAFR 207
      208 ERYRNIDVLLIDDI 221

QY      72 HRYREANIIMLDDI 85
      208 ERYRNIDVLLIDDI 221

```

Search completed: December 15, 2003, 15:19:02
 Job time : 16.0165 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 9.24897 Seconds
(without alignments)
1705.235 Million cell updates/sec

Title: US-09-689-952-16

Perfect score: 851

Sequence: 1 AADDICTAITNGEQVKGGLYL.....BRVKSLSPTPYFLSGENPRNN 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	851	100.0	306	2 A89952	primosomal protein
2	427.5	50.2	311	1 IQB544	primosome componen
3	423.5	49.8	307	2 AB1632	primosome componen
4	420.5	49.4	311	2 H84042	primosome componen
5	418.5	49.2	307	2 AH1269	primosome componen
6	347	40.8	298	2 B98066	primosome componen
7	346	40.7	298	2 D95199	primosomal protein
8	330	38.8	293	2 H86718	primosomal protein
9	166	19.5	209	2 S77882	dnab protein homol
10	140.5	16.5	282	2 D97138	DNA replication pr
11	140	16.5	316	2 B82807	conserved hypotet
12	129.5	15.2	235	2 E70378	DNA replication pr
13	127.5	15.0	313	2 D69945	phage-related prot
14	127	14.9	266	2 I40411	PBSX prophage ORF
15	125.5	14.7	440	2 D81415	chromosomal replic
16	114	13.4	265	1 BVECTT	1stb protein - Esc
17	111	13.0	350	2 B97735	probable atpase n2
18	109	12.8	316	2 T10436	probable transpos
19	105.5	12.4	248	2 C85626	probable DNA repli
20	105.5	12.4	248	2 B90763	probable DNA repli
21	105	12.3	246	2 C90913	probable DNA repli
22	104.5	12.3	248	1 C64886	DNA replication pr
23	104.5	12.3	248	2 A85472	probable DNA repli
24	104.5	12.3	248	2 A99872	probable DNA repli
25	104.5	12.3	287	2 AD1888	similarity Scap
26	104.5	12.3	287	2 AG1651	similarity Staph
27	104	12.2	520	2 H69125	hypothetical prote
28	103.5	12.2	248	2 AC0738	probable DNA repli
29	103.5	12.2	263	2 T47074	hypothetical prote

30	103.5	12.2	726	2 D72613	probable transito
31	103	12.1	454	2 D84931	chromosomal replic
32	102.5	12.0	245	2 D84932	DNA replication pr
33	102.5	12.0	432	2 S62461	probable AAA-famil
34	102.5	12.0	783	2 T18421	hypothetical prote
35	101.5	11.9	733	2 H69411	cell division cont
36	101	11.9	251	2 S23889	hypothetical prote
37	101	11.9	294	2 C66802	hypothetical prote
38	101	11.9	699	2 B72479	probable transito
39	100.5	11.8	357	2 D71732	probable atpase n2
40	100.5	11.8	811	2 B69512	cell division cont
41	100	11.8	291	2 A86681	hypothetical prote
42	100	11.8	307	2 G64239	Holliday junction
43	100	11.8	462	2 AG0497	chromosomal replic
44	100	11.8	656	2 A41870	dnab protein - Str
45	99	11.6	466	1 IQB5V	DNA replication initia

ALIGNMENTS

RESULT 1
A89952
primosomal protein [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C/Accession: A89952
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Ito, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; MUID:21311952; PMID:11418146
A/Accession: A89952
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-306 <KUR>
A/Cross-references: GB:BA000018; PID:g13701480; PIDN:BAB42774.1; GSPDB:GN00149
A/Experimental source: strain N315
C/genetics:
A/gene: dnaI
C/superfamily: 44K dnaa protein homolog

Query Match 100.0%; Score 851; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 5.4e-68;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AADDICTAITNGEQVKGGLYGPFGTKSFLIGAIANOLSKKVRSTIYLPFRTTLKG 60
DB 143 AADDICTAITNGEQVKGGLYGPFGTKSFLIGAIANOLSKKVRSTIYLPFRTTLKG 202

QY 61 GKQDSFEKKLRHREANILMDIGAEVTPWVDEVIQPLIHYRMVHELPTFPSSND 120
DB 203 GKQDSFEKKLRHREANILMDIGAEVTPWVDEVIQPLIHYRMVHELPTFPSSND 262

QY 121 YSELHHLAMTRDGEKTKAARIIRVKSLSPTPYFLSGENPRNN 164
DB 263 YSELHHLAMTRDGEKTKAARIIRVKSLSPTPYFLSGENPRNN 306

RESULT 2
IQB544
primosome component (helicase loader) dnaI - Bacillus subtilis
N/Alternate names: dna protein homolog, 44K, hypothetical protein Y (dnaB 3' region)
C/Species: Bacillus subtilis
C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 19-Jan-2001
C/Accession: B24720; C26580; F63617
R/Ogasawara, N.; Moriya, S.; Moriya, P. G.; Yoshikawa, H.
Nucleic Acids Res. 14, 9989-9999, 1986
A/Title: Nucleotide sequence and organization of dnaB gene and neighbouring genes on the
A/Reference number: A93650; MUID:87117549; PMID:3027671
A/Accession: B24720
A/Molecule type: DNA

A/Residues: 1-311 <OGA>
 A/Cross-references: GB:X04963; NID:G39880; PIDN:CAA2863.1; PID:G39881
 R/Hoshino, T.; McKenzie, T.; Schmidt, S.; Tanaka, T.; Sueoka, N.
 Proc. Natl. Acad. Sci. U.S.A. 84, 653-657, 1987
 A/Title: Nucleotide sequence of *Bacillus subtilis* dnaB: a gene essential for DNA replication
 A/Reference number: A94709; MUID:87118226; PMID:3027697
 A/Accession: C26580
 A/Molecule type: DNA
 A/Residues: 1-18, 'N', 20-23, 'T', 25-206 <HOS>
 R/Kunet, F.; Ogaawara, N.; Moszer, I.; Albertini, A.M.; Allont, G.; Azevedo, V.; Berth
 C./Bron, S.; Brouillet, S.; Bruchet, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cinc
 A./Ehlich, S.D.; Emmerston, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gallie
 tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
 Koetter, P.; Koningsberg, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A/Authors: Lander, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Toesto, V.; Uchiyama,
 T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yaeumoto, K.; Yata, K.; Yoshida, K
 A/Authors: Yoshikawa, H.F.; Zmistein, E.; Yoshikawa, H.; Danchin, A.
 A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A/Reference number: A69580; MUID:98044033; PMID:9384377
 A/Accession: F69617
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-311 <KUN>
 A/Cross-references: GB:Z99118; GB:AL009126; NID:G2635200; PIDN:CAB14858.1; PID:G2635363
 A/Experimental source: strain 168
 C/Genetics:
 A/Gene: dnaI
 C/Superfamily: 44k dnaa protein homolog
 C/Keywords: ATP; nucleotide binding; P-loop
 F:168-175/Region: nucleotide-binding motif A (P-loop)
 Query Match 50.2%; Score 427.5; DB 1; Length 311;
 Best Local Similarity 55.7%; Pred. No. 1.9e-30;
 Matches 83; Conservative 26; Mismatches 39; Indels 1; Gaps 1;
 QY 16 KGLVYGPFGTSGFLIGALANOLSKKVRSTIYLPDEFIRTLKGFQDGFSEKRLHVR 75
 DB 162 KGLVYKFGVGTGFMALANLAEKESMSIYVDEFEFRELKNSIQDQTLERKLNWVK 221
 QY 76 EANIMLDDIGAEVYTPWVDEVIQPLHYRMVHELTPFFSSNPDYSELEHILAMTRDGE 135
 DB 222 TTPVLMDDIGAESMTSWVRDEVIGTLQHRMSQQLPTFFSSNFPDELKHHFTYSQGRG 281
 QY 136 -EKTAKARIIRVKSITPYFLSGENFRN 163
 DB 282 KEVKAKARIIRVKSITPYFLSGENFRN 310
 RESULT 3
 ABI632
 primosome component (helicase loader) DnaI [imported] - *Listeria innocua* (strain Clp112
 C/Species: *Listeria innocua*
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C/Accession: ABI632
 R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Matounam, A.; Ma
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Matounam, A.; Ma
 A/Title: Comparative genomics of *Listeria* species.
 A/Reference number: AB1077; MUID:21537279; PMID:11679669
 A/Accession: ABI632
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-307 <GUA>
 A/Cross-references: GB:AL592022; PIDN:CAC96826.1; PID:G16414082; GSPDB:GN00178

A/Experimental source: strain Clp11262
 C/Genetics:
 A/Gene: dnaI
 C/Superfamily: 44k dnaa protein homolog
 Query Match 49.8%; Score 423.5; DB 2; Length 307;
 Best Local Similarity 52.3%; Pred. No. 4.3e-30;
 Matches 80; Conservative 32; Mismatches 40; Indels 1; Gaps 1;
 QY 11 NGRVYGVLYGPFGTSGFLIGALANOLSKKVRSTIYLPDEFIRTLKGFQDGFSEKRLHVR 70
 DB 154 NGRVYGVLYGPFGTSGFLIGALANOLSKKVRSTIYLPDEFIRTLKGFQDGFSEKRLHVR 213
 QY 71 LHRVREANIMLDDIGAEVYTPWVDEVIQPLHYRMVHELTPFFSSNPDYSELEHILAM 130
 DB 214 IQPARKTEVLMDDIGAESMTSWVRDEVIGTLQHRMSQQLPTFFSSNFPDELKHHFTYSQGRG 273
 QY 131 TRDG-EKTAKARIIRVKSITPYFLSGENFRN 162
 DB 274 AQNGTEKTLAKARIIRVKSITPYFLSGENFRN 306
 RESULT 4
 H84042
 primosome component (helicase loader) DnaI [imported] - *Bacillus halodurans* (strain C-12;
 C/Species: *Bacillus halodurans*
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C/Accession: H84042
 R/Takaki, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maui, N.; Fuji, F.; Hiran
 Nucleic Acids Res. 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: H84042
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-311 <STO>
 A/Cross-references: GB:AP001517; GB:BA000004; NID:G10175500; PIDN:BA00863.1; GSPDB:GN001
 A/Experimental source: strain C-125
 C/Genetics:
 A/Gene: dnaI
 C/Superfamily: 44k dnaa protein homolog
 Query Match 49.4%; Score 420.5; DB 2; Length 311;
 Best Local Similarity 50.0%; Pred. No. 8e-30;
 Matches 77; Conservative 33; Mismatches 43; Indels 1; Gaps 1;
 QY 12 GEGVYGVLYGPFGTSGFLIGALANOLSKKVRSTIYLPDEFIRTLKGFQDGFSEKRLHVR 71
 DB 156 GEGVYGVLYGPFGTSGFLIGALANOLSKKVRSTIYLPDEFIRTLKGFQDGFSEKRLHVR 215
 QY 72 HRYREANIMLDDIGAEVYTPWVDEVIQPLHYRMVHELTPFFSSNPDYSELEHILAM 131
 DB 216 DRYKNAQVLFDDIGAEVYTPWVDEVIQPLHYRMVHELTPFFSSNPDYSELEHILAM 275
 QY 132 -RDSEKTKARIIRVKSITPYFLSGENFRN 164
 DB 276 DKSGTEKTLAKARIIRVKSITPYFLSGENFRN 309
 RESULT 5
 AH1269
 primosome component (helicase loader) DnaI [imported] - *Listeria monocytogenes* (strain EC
 C/Species: *Listeria monocytogenes*
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C/Accession: AH1269
 R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Matounam, A.; Ma
 A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Matounam, A.; Ma
 A/Title: Comparative genomics of *Listeria* species.
 A/Reference number: AB1077; MUID:21537279; PMID:11679669

N:Alternate names: protein MC072
 C:Species: Mycoplasma capricolum
 C>Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 07-Dec-1999
 C/Accession: S77882
 R:Bork, P.; Ouzounis, C.; Casari, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.;
 Mol. Microbiol. 16, 955-967, 1995
 A>Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its phyletic
 A/Reference number: S77739; MUID:9605641; PMID:7476192
 A/Accession: S77882
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-209 <BCR>
 A/Cross-references: EMBL:Z33058; NID:G514450; PID:CAA83732.1; PID:G530419
 A/Experimental source: ATCC 27343
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
 C/Genetics:
 A/Genetic code: SGC3

Query Match 19.5%; Score 166; DB 2; Length 209;
 Best Local Similarity 35.9%; Pred. No. 1.8e-07;
 Matches 37; Conservative 21; Mismatches 39; Indels 6; Gaps 3;
 QY 16 KGLVLPFGTSGFLLGAIANOLSKKVRSTIY--LPEFIRTLKGGFGSGFE--KTL 71
 DB 100 KGVLYGHSGIGKTYMKVLANLASK--NKTVIFSTRSLIDRLKESFNSSEINSLM 157
 QY 72 HRREANIIMLDDIGAEVTPWVRDEVIGPLHYRMVHELPTF 114
 DB 158 KKIKTVDPLFLDDIGENLSLWARDPLFEVLYNRMENOKATP 200

RESULT 10

D97138
 DNA replication protein DnaC [imported] - Clostridium acetobutylicum
 C/Species: Clostridium acetobutylicum
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C/Accession: D97138
 R:Jolling, J.; Bretton, G.; Omeikhenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koehn, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4623-4838, 2001
 A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: D97138
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-282 <KRR>
 A/Cross-references: GB:AE001437; PID:AAK79895.1; PID:G15024913; GSPDB:GN00168
 A/Experimental source: Clostridium acetobutylicum ATCC824
 C/Genetics:
 A/Genetic code: CAC1933

Query Match 16.5%; Score 140.5; DB 2; Length 282;
 Best Local Similarity 23.5%; Pred. No. 4.7e-05;
 Matches 36; Conservative 40; Mismatches 68; Indels 9; Gaps 3;
 QY 13 EYKVGVLVGPFGTSGFLLGAIANOLSKKVRSTIY--LPEFIRTLKGGFGSGFEKTL 72
 DB 120 QQGLGLIYSGAGNKYTLASAIANELLKOYIPVVCVINGLSRIQKTYNSNGKEASD 179
 QY 73 RVR---EANTIMLDDIGAEVTPWVRDEVIGPLHYRMVHELPTFSSNFDYSELEHHLA 129
 DB 180 IISFIMADLIIDDLCTEKSEMSR-SMTYINVDSSYRSKLPILITNSLEINPSKRGHV 238
 QY 130 MTRDGEKTKAARIERVKSLSTPYLGENFR 162
 DB 239 IADQYHERTES-----RIFEMCTPVENTSKSIR 266

RESULT 11

B82907
 conserved hypothetical ATP/GTP-binding protein UU307 [imported] - Ureaplasma urealyticum
 C/Species: Ureaplasma urealyticum
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C/Accession: B82907
 R:Glase, J.I.; Ieffkowitz, E.J.; Glase, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
 Submitted to Genbank, February 2000
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mini
 A/Reference number: A82870
 A/Accession: B82907
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1316 <GLA>
 A/Cross-references: GB:AE002128; GB:AF222894; NID:G6699279; PID:AAF30716.1; GSPDB:GN001;
 A/Experimental source: serovar 3; biovar 1
 C/Genetics:
 A/Genetic code: SGC3

Query Match 16.5%; Score 140; DB 2; Length 316;
 Best Local Similarity 25.5%; Pred. No. 6e-05;
 Matches 39; Conservative 28; Mismatches 68; Indels 18; Gaps 3;
 QY 11 NGBQVGLVGPFGTSGFLLGAIANOLSKKVRSTIY--LPEFIRTLKGGFGK----- 63
 DB 151 NNQPKSARLYGDFSVGKSIITQATYNTI--SLKTNLKIAYT-----TLNDLFKNVYQFEN 204
 QY 64 -----DSEFEKTLHVRREANIIMLDDIGAEVTPWVRDEVIGPLHYRMVHELPTFSSN 118
 DB 205 YKOTSDLVINELVNLNDVLDVIDPFSSVNLNWSISITLWPIENRLKSTRTQITFISN 264
 QY 119 FDYSLEHHLAMTRDGEKTKAARIERYKSL 151
 DB 265 FSIEQLINSTKNTNIEOKTKLRFNRLEYLT 297

RESULT 12

E70378
 DNA replication protein DnaC - Aquifex aeolicus
 C/Species: Aquifex aeolicus
 C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C/Accession: E70378
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
 V.
 Nature 392, 353-358, 1998
 A>Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A/Reference number: A70300; MUID:98196666; PMID:9537320
 A/Accession: E70378
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-235 <NAQ>
 A/Cross-references: GB:AE000713; NID:G2983424; PID:AACT013.1; PID:G2983431; GB:AE00065;
 A/Experimental source: strain VFS
 C/Genetics:
 A/Genetic code: dnaC

Query Match 15.2%; Score 129.5; DB 2; Length 235;
 Best Local Similarity 26.5%; Pred. No. 0.00056;
 Matches 45; Conservative 27; Mismatches 61; Indels 37; Gaps 5;
 QY 11 NGBQVGLVGPFGTSGFLLGAIANOLSKK--VRSTIY--LPEFIRTLKGGFGKSGFEK 69
 DB 75 NPEBGKLTFFVSGPGKTHLVAATLKAIEKKGKNGYFDDTKDLFRKHLMDGKDK 134
 QY 70 KLRHVRANIIMLDDIGAEVTPWVRDEVIGPLHYRMVHELPTFSSNFDYSELEHHLA 129
 DB 135 FLKTVANSFVLDVDDIGSERLSDWQR-ELISYITITRYNNKSLTITTY-----S 184
 QY 130 MTRDGEKTK-----AARIERY-----KSLSTP 153
 DB 185 LQREESSVRISADLASRLGENVVSKEYEMNELLVKSGDLRKSKGLSTP 234

RESULT 13

D69945
 phage-related protein homolog yqam - Bacillus subtilis
 C/Species: Bacillus subtilis

```

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: C69731
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-238, 'K', 240-266 <KUN>
A:Cross-references: GB:599110; GB:AL009126; NID:G2633472; PIDN:CAB13110.1; PID:e1183273
C:Experimental source: strain 168
C:Genetics:
A:Gene: xkdc

Query Match      14.9%; Score 127; DB 2; Length 266;
Best Local Similarity 31.0%; Pred. No. 0.00069;
Matches 39; Conservative 25; Mismatches 44; Indels 18; Gaps 7;

Qy      6 CTA--ITNGQVK-----GLYVGPFGTGKSPILGAIANOLKSKRVSTIIYLP--EFIR 56
      ||-:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      101 CTKEVYADVADYQIDCKRNGSIALLGQPSGKTHLLTAANEL--MTCVCPVPIYFPFVGFT 159
      ||-:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy      .57 TLKGGFKDGSFEKKLHFVRERANITIMLDI-----GAEVTPWVADDEVIGPLHRYMVEL 111
      ||-:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      160 DLKNDF--ALLKATLNMRKQADVLFIDLLFKFPVNGKRPATDMLQGYVS--VLNRYILNHK 216
      ||-:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy      112 PTFSS 117
      ||-:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 15

Chromosomal replication initiator protein Cj0001 [imported] - Campylobacter jejuni (str) NC_008695.1
C.Species: Campylobacter jejuni
C.Date: 31-Mar-2000 #sequence_revision: 31-Mar-2000 #text_change: 03-Jun-2002
C.Accession: D81415
R.Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Whitehead, P.A.; Barrell, A.G.; Paulsen, O.; Karpman, D.; Davies, R.J.; Holt, K.E.; Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chillingworth, T.; Whitehead, P.A.; Barrell, A.G.; Paulsen, O.; Karpman, D.; Davies, R.J.
C.W., Quail, M., Rajandream, M.A., Rutherford, K.M., VanVleet, A., Whitchaed, S., Barre, N., Quail, M., Rajandream, M.A., Rutherford, K.M., VanVleet, A., Whitchaed, S., Barre, N.
A.Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hlyN
A.Reference number: AB12502; PMID:20150912; PMID:10688204
A.Accession: D81415
A.Status: Preliminary

A;Cross-references: GB:AL1139074; GB:AL1111166; NID:96967505; PIDN:CA872494.1; PID:96967505
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: dnaA; Cj0001

Query Match	14.7%	Score 125.5	DB 2	Length 440
Best Local Similarity	26.8%	Pred. No. 0.0017		
Matches	41	Conservative	30	Mismatches 51
				Indels 31
				Gaps 7
6	CTAATNGEQVKGLY--- <td>59</td> <td></td> <td></td>	59		

```

Db      123  CKIAIHDCKKCKLYNPIFFVVGPTGLCKTHLLOAVGN--ASLEMGKVIYATSENINDFT 180
Oy      60  GGPFQDSFEKKLHRVREANTIMLDDIGAEETVPWVRDEVIGPLLHYRMVHELPTPFSSNF 119
Db      181  SNLKNGSLDFHEKRYRRCVDVLIDVQFLCKTDKIGEE-----F---F 221
Oy      120  DYSELEHH--LAWTRDGECKTKAARIIEVKVS 149
Db      222  IFNEIKNNDOIIMTSDNPPNMLKG-ITERLKVS 253

Search completed: December 15, 2003, 15:16:49
Job time : 9.24897 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 5.18645 Seconds
(without alignments)
1486.453 Million cell updates/sec

Title: US-09-689-952-16

Perfect score: 851

Sequence: 1 AADICTAITNGEQVKGVL.....ERVKSLSTPYFLSGENFRNN 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	427.5	50.2	311 1	DNA1_BACSU
2	127.5	15.0	313 1	YQAM_BACSU
3	127	14.9	266 1	XKDC_BACSU
4	125.5	14.7	440 1	DNA1_CAMJE
5	114	13.4	265 1	ISTB_PSBAS
6	106.5	12.5	246 1	DNA1_BUCBP
7	106	12.5	457 1	DNA1_BUCBP
8	105.5	12.4	229 1	ISTB_BURCE
9	104.5	12.3	248 1	YDVA_ECOLI
10	103.5	12.2	263 1	Y4BW_RHISN
11	103	12.1	454 1	DNA1_BUCAI
12	102.5	12.0	246 1	SKDI_SCHPO
13	102.5	12.0	432 1	ISTB_SCHPO
14	101	11.9	251 1	ISTB_BACST
15	100	11.8	307 1	RUVB_MYCGE
16	100	11.8	462 1	DNA1_YERPE
17	100	11.8	464 1	DNA1_SERMA
18	100	11.8	643 1	DNA1_STRRE
19	100	11.8	656 1	DNA1_STRRE
20	99.5	11.7	231 1	ISTB_PSEFL
21	99	11.6	466 1	DNA1_PROMT
22	98	11.5	454 1	DNA1_BUCAP
23	98	11.5	466 1	DNA1_SALTI
24	98	11.5	466 1	DNA1_SALTY
25	98	11.5	467 1	DNA1_ECO57
26	98	11.5	467 1	DNA1_ECOLI
27	97.5	11.5	451 1	DNA1_PASMT
28	97	11.4	245 1	DNA1_ECOLI
29	97	11.4	263 1	ISTB_BACFR
30	96	11.3	410 1	PSMR_METTH
31	95.5	11.2	340 1	RFC5_HUMAN
32	95.5	11.2	453 1	DNA1_STRPN
33	95	11.2	468 1	DNA1_VIRBU

34	95	11.2	624 1	DNA1_STRCH	09zh75 streptomyc
35	95	11.2	758 1	SC18_YEAST	P18759 saccharomyc
36	94.5	11.1	399 1	DNA1_AQUAE	O66559 aquifex aco
37	94.5	11.1	794 1	SC18_CANAL	P34732 candida ald
38	94	11.0	398 1	PSMR_ARCFU	O28303 archaeoglob
39	94	11.0	436 1	PSMR_METKA	O8x03 methanopyru
40	94	11.0	451 1	DNA1_LISNO	O92lv2 listeria in
41	94	11.0	451 1	DNA1_LISNO	O8yav2 listeria in
42	94	11.0	484 1	DNA1_ZYMO	O98693 zymomonas m
43	93	10.9	437 1	DNA1_YEAST	P52917 saccharomyc
44	93	10.9	463 1	DNA1_RICPR	O59758 rickettsia
45	92	10.8	250 1	ISTB_BACTB	O99338 bacillus th

ALIGNMENTS

RESULT 1	ID	DNA1_BACSU	STANDARD	PRT	311 AA.
AC	P06567				
DT	01-JAN-1988	(Rel. 06, Created)			
DT	01-JAN-1988	(Rel. 06, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	Primosomal protein dna1.				
GN	DNA1				
OS	Bacillus subtilis.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxID=1423;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87117549; PubMed=3027671;				
RA	Ogasawara N., Moriya S., Mazza P.G., Yoshikawa H.;				
RT	"Nucleotide sequence and organization of dnaB gene and neighbouring				
RT	genes on the Bacillus subtilis chromosome."				
RL	Nucleic Acids Res. 14:9989-9999 (1986).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=168;				
RX	MEDLINE=97124191; PubMed=8969504;				
RA	Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.;				
RT	Sanders J., Emerson P.T., Harwood C.R.;				
RT	"The dnaB-phaA (256 degrees-240 degrees) region of the Bacillus				
RT	subtilis chromosome containing genes responsible for stress				
RT	responses, the utilization of plant cell walls and primary				
RL	metabolism."				
RN	Microbiology 142:3067-3078 (1996).				
RP	[3]				
RC	SEQUENCE FROM N.A.				
RX	STRAIN=168;				
RA	Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;				
RT	"Sequencing and functional annotation of the Bacillus subtilis genes				
RT	in the 200 kb rmb-dnaB region."				
RL	Microbiology 143:3431-3441 (1997).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=168;				
RX	MEDLINE=98044033; PubMed=9384377;				
RA	Kuntz F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.;				
RT	Azeredo V., Barreiro M.G., Bessieres P., Bolotin A., Borchert S.;				
RT	Borrius R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.;				
RA	Bouilliet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.;				
RA	Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.;				
RA	Dentant F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.;				
RA	Eutlian K.D., Etrington J., Fabret C., Ferrari B., Foulger D.;				
RA	Fritz C., Fujita P., Fujita Y., Fuma S., Galizzi A., Galleron N.;				
RA	Ghim S.Y., Glaser P., Goffeau A., Colightly E.J., Grandi G.;				
RA	Gutseppe G., Guy B.J., Haga K., Haeche J., Harwood C.R., Henaut A.;				
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.;				
RA	Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.;				
RA	Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.;				
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.;				

RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moserl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
RA Prescan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccanti E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosafo V., Uchiyama S., Vandenbol M., Vannier F., Vassaroletti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,"
RL Nature 390:249-256(1997).
RN [5]
RP SEQUENCE OF 1-206 FROM N.A.
RX MEDLINE=87118226; PubMed=3027697.
RA Hoshino T., McKenzie T., Schmidt S., Tanaka T., Sueoka N.,
RT "Nucleotide sequence of *Bacillus subtilis* dnaB: a gene essential for
RT DNA replication initiation and membrane attachment,"
RL Proc. Natl. Acad. Sci. U.S.A. 84:653-657(1987).
RN [6]
RP SIMILARITY TO DNA.
RX MEDLINE=92195821; PubMed=1549481,
RA Koonin E.V.,
RT "Archaeobacterial virus SSV1 encodes a putative DnaA-like protein,"
RL Nucleic Acids Res. 20:1143-1143(1992).
RN [7]
RP IDENTIFICATION.
RX MEDLINE=95291463; PubMed=7773414,
RA Brand C., Ehrlich S.D.,
RT "The *Bacillus subtilis* dnaI gene is part of the dnaB operon,"
RL Microbiology 141:1199-1200(1995).
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: SOME, TO DNA FROM VARIOUS BACTERIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X04963; CAA28633.1; -
DR EMBL: 275208; CAA99605.1; -
DR EMBL: AF008220; AAC00359.1; -
DR EMBL: Z99118; CAB14858.1; -
DR EMBL: M15183; AAA22405.1; -
DR PIR: B24720; I0BS44.
DR Subtilisin; BGI0359; dnaI.
DR Interpro: IPR003593; AAA_ATPase.
DR SMART: SMO0382; AAA: 1
KW Primosome; DNA replication; ATP-binding; Complete proteome.
FT NP_BIND 168 175 ATP (PROBABLY).
FT CONFLICT 19 19 K -> N (IN REF. 5).
FT CONFLICT 24 24 M -> T (IN REF. 5).
SQ SEQUENCE 311 AA; 36114 MW; A86FC94AB6841264 CRC64;
Query Match 50.2%; Score 427.5; DB 1; Length 311;
Best Local Similarity 55.7%; Pred. No. 1.8e-31;
Matches 83; Conservative 26; Mismatches 39; Indels 1; Gaps 1;
QY KGLVYGPFGKSPFIATANOLKSKVSTIYLEPFRTLLKGFKDSPEKKLRVR 75
DB 162 KGLVYKFGVQKTFMLAALANLAEKYSMTYVEFRELKNSLDOTLEKLMVK 221
QY 76 EANILMDIDIGAEVTPWVDEVIQPLHYRMVHELPFPSSNEDVSELEHLLAMTRDGE 135
DB 222 TTPVLMDDIGASMTSWVDEVIQVTLQHRMSQQLPFPSSNSNPDLKHHFVYSGRGE 281

QY 136 -EXTRAKTIERVKSLSTPYFLSGENFRN 163
DB 282 KESEVKARLMERTLYLAAPRIJDGENRRH 310
RESULT 2
ID YQAM_BACSU STANDARD, PRT, 313 AA.
AC P45910;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yqam.
GN YQAM.
OS *Bacillus subtilis*.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_Taxid=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=95219086; PubMed=7704261;
RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.,
RT "Complete nucleotide sequence of a skin element excised by DNA
RT rearrangement during sporulation in *Bacillus subtilis*,"
RL Microbiology 141:323-327(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=97124195; PubMed=8969508;
RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
RA Kobayashi Y.,
RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
RT the *Bacillus subtilis* genome containing the skin element and many
RT sporulation genes,"
RL Microbiology 142:3103-3111(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
RA Azevedo V., Betero M.G., Bessieres P., Bolocin A., Borchert S.,
RA Borriss R., Bourster L., Brans A., Braun M., Britnell S.C., Bron S.,
RA Brouillet S., Bruch C.V., Caldwell I.F., Cummings N.J., Carter N.M.,
RA Choi S.K., Codani J.U., Comercon I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Duberthof A., Ehrlich S.D., Emmerson P.T.,
RA Denton K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Golligly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Iaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moserl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
RA Prescan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccanti E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosafo V., Uchiyama S., Vandenbol M., Vannier F., Vassaroletti A.,
RA Viari A., Wambut R., Wedler E., Wedler H., Weltzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,"
RL Nature 390:249-256(1997).
RN [4]
RP IDENTIFICATION.

RA MEDLINE=96084975; PubMed=7489895;
 RA Medigue C., Moszer I., Vlati A., Danchin A.;
 RA "Analysis of a Bacillus subtilis genome fragment using a co-operative
 RT computer system prototype";
 RL Gene 165:GC37-GC51(1995).
 CC -1- SIMILARITY: TO B.SUBTILIS Y0XC AND T.HYDROXYBENTYL HEMOLYSIN
 TLVA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D32216; BAA06927.1; -
 DR EMBL; D84432; BAA12388.1; -
 DR EMBL; Z99117; CAB14567.1; -
 DR PIR; D69945; D69945.
 DR Subtilisin; BG11264; ygaM.
 DR InterPro; IPR003593; AAA_ATPase.
 DR SMART; SM00382; AAA; 1.
 DR Hypothetical protein; Complete proteome.
 SQ SEQUENCE 313 AA; 36138 MW; 09208795E310D73A CRC64;
 Query Match 15.0%; Score 127.5; DB 1; Length 313;
 Best Local Similarity 33.1%; Pred. No. 0.00023;
 Matches 40; Conservative 19; Mismatches 49; Indels 13; Gaps 6;
 QY 12 GEQVKKLYIGPFGTGSFLGAIANQL-KSKYVRSTIYLP-EFIRTLKGFGKGSPE 68
 Db 161 GERQNSIALLGQSGSKTHLTAMNLIKKSIVH-CMFYFVEGMGLKANF-DNIE 216
 QY 69 KKLHVRVANIIMLDI-----GAEVTPVRNDEVIGPLHYRMVHELPTFGSSNDYSE 123
 Db 217 AKLDARKEVLEFIDLFKPIGQPRATPW-QVEQLQSVLVNRYVLMHKKPLISSSELTIDE 275
 QY 124 L 124
 Db 276 I 276
 RESULT 3
 XKDC_BACSU STANDARD; PRT; 266 AA.
 ID XKDC_BACSU STANDARD; PRT; 266 AA.
 AC P39782;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phage-like element PBSX protein XKDC.
 GN XKDC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=1423;
 RX NCB1_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / SO113;
 RC MEDLINE=94364963; PubMed=8083174;
 RA McDowell G.E., Wood H., Devine K.M., McConnell D.J.;
 RT "Genetic control of bacterial suicide: regulation of the induction of
 RT PBSX in Bacillus subtilis";
 RT J. Bacteriol. 176:5820-5830(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RC Krogh S., O'Reilly M., Nolan N., Devine K.M.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RC MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boris R., Bourcier L., Brans A., Braun M., Briganti S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Comerion I.F., Cummings N.J., Daniel R.A.,
 RA Denton J.F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giesepf G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianhard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapides A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestel D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Puig P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan B., Schleich S., Schroeter R., Scoffone P.,
 RA Sekiguchi J., Sekowska A., Serró S.J., Serró P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi Y., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandendol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Welzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis";
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTIONAL ANTITERMINATOR.
 CC -1- SIMILARITY: STRONG, TO B.SUBTILIS Y0AM.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z34287; CAA64044.1; -
 DR EMBL; Z70177; CAA94054.1; -
 DR EMBL; Z99110; CAB13110.1; -
 DR PIR; I40411; I40411.
 DR Subtilisin; BG10996; xkdc.
 DR InterPro; IPR003593; AAA_ATPase.
 DR SMART; SM00382; AAA; 1.
 KW ATP-binding; Complete proteome.
 FT NP BIND 124 131 ARP (POTENTIAL).
 FT CONFICT 239 239 K -> R (IN REF. 1).
 SQ SEQUENCE 266 AA; 30534 MW; 9929C991E9D655AA CRC64;
 Query Match 14.9%; Score 127; DB 1; Length 266;
 Best Local Similarity 31.0%; Pred. No. 0.00021;
 Matches 39; Conservative 25; Mismatches 44; Indels 16; Gaps 7;
 QY 6 CTA-ITNGEYK-----GLVYIGPFGTGSFLGAIANQLSKYVRSTIYLP-EFIR 56
 Db 101 CKEVYADVEIQKDKRNSIALLGQSGSKTHLTAAANL-MRTCYVVIYFPEFGFT 159
 QY 57 TLKGFKGSFKKLHVRVANIIMLDI-----GAEVTPVRNDEVIGPLHYRMVHEL 111
 Db 160 DLKNDP--ALTEAKLNRKQADVLFDLFKPVNGKPRATPWQLEQMS-VLNYRYLHKK 216
 QY 112 PTFSS 117
 Db 217 PILLS 222
 RESULT 4
 DNA_CAMJ DNA_CAMJ STANDARD; PRT; 440 AA.

AC 09PJ001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chromosomal replication initiator protein dnaA.
 GN DnaA OR C30001.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Campylobacteraceae; Campylobacter.
 OK NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 Basham D., Chillingworth T., Davies K.M., Felwell T., Holtroyd S.,
 Jaseis K., Kariyeh A.V., Moule S., Pallen M.J., Penn C.W.,
 Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 Whitehead S., Barrall B.G.;
 RA Whitehead S., Barrall B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 reveals hypervariable sequences."
 RL Nature 403:665-668(2000).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
 OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
 IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
 (DnaA BOX): 5'-TTATTC(C/A)A(C/A)A-3'. DnaA BINDS TO ATP AND TO
 CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the dnaA family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AL139074; CAB72494.1; -
 DR PIR: D81415; D81415.
 DR HAMAP: MF_003771; -1.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001957; Bac_DnaA.
 DR Pfam: PF00308; bac_dnaA; 1.
 DR PRINTS: PRO0051; DnaA.
 DR SMART: SM00382; AAA; 1.
 DR TIGRFA: TIGR00362; DnaA; 1.
 DR PROSITE: PS01008; DnaA; FALSE_NEG.
 KW DnaA replication; DNA-binding; ATP-binding; Complete proteome.
 FT NP BIND 143 150 ATP (POTENTIAL).
 SQ SEQUENCE 440 AA; 49686 MW; A3FAB90CCED06B07 CRC64;
 Query Match 14.7%; Score 125.5; DB 1; Length 440;
 Best Local Similarity 26.8%; Pred. No. 0.00052;
 Matches 41; Conservative 30; Mismatches 51; Indels 31; Gaps 7;
 QY 6 CTATNGEYQKGY-----LYGPGTGSFLLGAIANDLKSXYRSTIYV--EPIRTLK 59
 DB 123 CKAIAHDKKQKJNPIYVGPGLGKTHLQAVGN--ASLENGKVIYATSNFINDFT 180
 QY 60 GGFKDGSEFKKLRVREANILMDIGAEVTPWVDEVIPLHYRMVHELPFFSSNF 119
 DB 181 SNKNGSLDGFHKRYRNCVDLIDVQFLGKTDKIOE-----FF---F 221
 QY 120 DYSELEHH--LAMTRDSEKTYAARIYRVS 149
 DB 222 IFNEIKNDQIIMTSDNPNMLKG-ITERLKS 253
 RESULT 5
 ID ISTB_PSEAE STANDARD; PRT; 265 AA.
 AC P15026;
 DT 01-APR-1990 (Rel. 14, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Insertion sequence IS21 putative ATP-binding protein.
 GN ISTB.
 OS Pseudomonas aeruginosa.
 OS Plasmid R68.45.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OK NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89218951; PubMed=2540414;
 RA Reimann C., Moore R., Little S., Savoy A., Willetts N.S., Haas D.;
 RT "Genetic structure, function and regulation of the transposable
 RT element IS21."
 RL Mol. Gen. Genet. 215:416-424(1989).
 RN [2]
 RP REVISION TO 283.
 RA Berger B.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE ISTB PROTEIN, ONE OF TWO PROTEINS EXPRESSED ONLY
 CC WHEN THERE IS A TANDM REPEAT OF THE IS21 INSERTION SEQUENCE, IS
 CC NECESSARY FOR THE TRANSPOSITION OF PLASMIDS WITH THAT TANDM
 CC REPEAT.
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING
 CC PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X14793; CAA32899.2; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR002611; IscB_ATPbind.
 DR Pfam: PF01695; IscB; 1.
 DR SMART: SM00382; AAA; 1.
 KW Plasmid; Transposable element; ATP-binding.
 FT NP BIND 104 111 ATP (POTENTIAL).
 SQ SEQUENCE 265 AA; 30528 MW; 904C7ADC583A12A9 CRC64;
 Query Match 13.4%; Score 114; DB 1; Length 265;
 Best Local Similarity 26.1%; Pred. No. 0.0032;
 Matches 42; Conservative 36; Mismatches 53; Indels 30; Gaps 8;
 QY 13 EOYKGLYLPFGTGSFLLGAIANDLKSXYRSTIYV--EPIRTLKGFPGSFEK 70
 DB 95 ERSENVILGPPGVGKTHL--AIALGVKAVDAGHRVLPDLRLATLTKAKQENRLERQ 152
 QY 71 LHRVREANILMDIGAEVTPWVDEVIPLHYRMVHELPFFSSNF--FDYSE 123
 DB 153 LQOLSVARVILIDELG--YLPNNRREA--SLPFLRLRRYRKASIIILSNKGFADMG 206
 QY 124 L--EHLAMTRDSEKTYAARIYRVSSTYPLSGENFR 162
 DB 207 MFGDHYL-----ATAILRLHLHSTLTKGESYR 236
 RESULT 6
 ID DNAC_BUCAP STANDARD; PRT; 246 AA.
 AC O8KA79;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA replication protein dnaC.
 GN DNAC OR BUSG022.
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.

OX NCBI_TaxID=98794;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22084549; PubMed=12089438;
 RA Tamas I., Klasson L., Candaeck B., Naeslund A.K., Eriksson A.-S.,
 Wernegreen J.U., Sandstrom J.P., Moran N.A., Andersson S.G.E.;
 RT "50 million years of genomic stasis in endosymbiotic bacteria.";
 RL Science 296:2376-2379(2002).
 CC -1- FUNCTION: THIS PROTEIN IS REQUIRED FOR CHROMOSOMAL REPLICATION. IT
 CC FORMS, IN CONCERT WITH DNAB PROTEIN AND OTHER PREPROMING PROTEINS
 CC DNAT, N', N', A PREPROMING PROTEIN COMPLEX ON THE SPECIFIC SITE
 CC OF THE TEMPLATE DNA RECOGNIZED BY PROTEIN N' (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DNAC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE014077; AAM67594.1; -
 KW DNA replication; Primosome; Complete proteome.
 FT SITE 69
 FT SEQUENCE 246 AA; 28497 MW; B7E2A9BC2ADB5D CRC64;
 SQ
 Query Match 12.5%; Score 106.5; DB 1; Length 246;
 Best Local Similarity 25.2%; Pred. No. 0.014;
 Matches 35; Conservative 29; Mismatches 60; Indels 15; Gaps 4;
 OY 13 EGVKGLYXGPGTGSFLGATANOLKSKKRVSTIYLPERRTLKGFQKSGS---FKK 69
 Db 97 ENIASFTSGRPTGNHLASALGNLILHGKSLILVTAADLSNNKGFPGSGTSNITEEN 156
 OY 70 KAHRYEANIIMLDIGAEVTPWVDEVYIGPLHYRMVHELPTFPSSNFDYSELEHHA 129
 Db 157 LHMNLSVDLIMIDELGQMTSEYER-VIINOIVDRSSSKSTGMLSLNDRGMGNLL- 214
 OY 130 MTRDGEKTKAARIIRVK 148
 Db 215 ---GB-----RVIDRMR 223
 RESULT 7
 DNAA_BUCBP STANDARD; PRT; 457 AA.
 ID DNAA_BUCBP
 AC P59567;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Chromosomal replication initiator protein dnnaA.
 GN DNAA OR BBP012.
 OS Buchnera aphidicola (subsp. Baizongia pistaciace).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 NCBI_TaxID=135842;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426901; PubMed=12522265;
 RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
 Bascolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
 Tames J., Viguera B., Latorre A., Valencia A., Moran F., Moya A.;
 RT "Reductive genome evolution in Buchnera aphidicola.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
 CC -1- FUNCTION: Plays an important role in the initiation and regulation
 CC of chromosomal replication. Binds to the origin of replication; it
 CC binds specifically double-stranded DNA at a 9 bp consensus (dnna
 CC box): 5'-TTATC(A/A)A(C/A)A-3'. Dnaa binds to ATP and to acidic
 CC phospholipids (By similarity).
 CC -1- SIMILARITY: Belongs to the dnna family.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE014016; AAO26756.1; -
 DR HAMAP; MF_00377; -; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS01008; DNAA; FALSE NEG.
 KW DNA replication; DNA-binding; ATP-binding; Complete proteome.
 FT NP BIND 162
 FT SEQUENCE 457 AA; 53365 MW; B2950813FADDA028 CRC64;
 SQ
 Query Match 12.5%; Score 106; DB 1; Length 457;
 Best Local Similarity 24.1%; Pred. No. 0.032;
 Matches 41; Conservative 30; Mismatches 71; Indels 28; Gaps 6;
 OY 18 LYLVPFGTGSFLGATANOLKSKKRVSTIYLPERRTLKGFQKSGFEKLRVR 75
 Db 158 LFLYKSGGLGKTHLHAAVANTLKYKNTIKIYINSENFQMTSLKNTIEFPKYR 217
 OY 76 EANIIMLDIGAEVTPWVDEV---IGPLHY-----RMVHELPTFPSSNFD 120
 Db 218 SVNTLIDIDIOFPAVKKSQSELEFHTNALNRNOQIIITSQFPQKHGIETRLKSR- 276
 OY 121 YSELEHHLAMTRDGEKTKAARI---ERVKSLSTPY---FUSGENFRNN 164
 Db 277 ---BCGLTIRIDPPDLMTRTKILIKSHIYDINLSYKVAFFIAKNLKSXN 322
 RESULT 8
 ISTR_BURCE STANDARD; PRT; 229 AA.
 ID ISTR_BURCE
 AC P55923;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Insertion sequence IS408 putative ATP-binding protein (ORF2).
 OS Burkholderia cepacia (Pseudomonas cepacia).
 OC Burkholderia; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia.
 NCBI_TaxID=292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 17616 / 249;
 RX MEDLINE=94302134; PubMed=7518087;
 RA Byrne A.M., Leslie T.G.;
 RT "Characteristics of IS401, a new member of the IS3 family implicated
 RT in plasmid rearrangements in Pseudomonas cepacia.";
 RL Plasmid 31:138-147(1994).
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING
 CC PROTEIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L09108; -; NOT ANNOTATED CDS.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR002611; IsetB_Atpbind.
 DR Pfam; PF01695; IsetB; 1.
 DR SMART; SM00382; AAA; 1.
 KW Transposable element; ATP-binding.
 FT NP BIND 105
 FT SEQUENCE 229 AA; 25929 MW; ACS619C729163E1 CRC64;
 SQ

Query Match 12.4%; Score 105.5; DB 1; Length 229;
 Best Local Similarity 26.8%; Pred. No. 0.016;
 Matches 39; Conservative 29; Mismatches 64; Indels 11; Gaps 5;

QY 6 CTAITNGEYQKGLVYPPFGTGSFIIIGAIANQLSKKVRSTIYLPEFIRTLKGGF 65
 DB 92 CDWRNNAQ---NLTLPPTGAGKWLACAFQOACROGFSFYVARLPEELKIHAGDG 148
 QY 66 SPEKKLRVREANILMDDIGAEVTPVWDEVGILLHYRMHTELPTFFSSNDYSELE 125
 DB 149 SFRRLLQMLADIVLLIDDLGOLDOAARNDLL-EVLDDR-VGTRSTVITSQI---PLE 203
 QY 126 HHLAMTRDGEKTKVARIIRV 147
 DB 204 HWHAWLQD---PTLADAILDLRL 222

RESULT 9
 YDAB_ECOLI STANDARD; PRT; 248 AA.

AC P75546;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Hypothetical protein ydab.
 GN YDAB OR B1360.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=9725157; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isano K., Itoh T.,
 RA Kasa H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakada S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Salto N.,
 RA Sempel G., Seki Y., Sivasubraman S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map.";
 RL DNA Res. 3:363-377(1996).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL: AE000233; AAC74442.1; -
 EMBL: D90774; BAI1958.1; -
 DR EIR; C64886; C64886.
 DR EcoGene; EG13364; ydab.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR002611; IetB_Atpbind.
 DR Pfam; PF01695; IetB_1.
 DR SMART; SM00382; AAA; 1.
 KW Hypothetical protein; Coiled coil; Complete proteome.
 FT DOMAIN 33 57 COILED COIL (POTENTIAL).
 SQ SEQUENCE 248 AA; 28128 MW; 3501DCB1E8512E01 CRC64;

Query Match 12.3%; Score 104.5; DB 1; Length 248;
 Best Local Similarity 27.2%; Pred. No. 0.021;
 Matches 40; Conservative 28; Mismatches 60; Indels 19; Gaps 6;

QY 2 ADDICTAINTGEYQKGLVYPPFGTGSFIIIGAIANQLSKKVRSTIYLPEFIRTLKGG 61
 DB 94 ADELMTGCTN-----FAFGKRGPTGKNHMAAIGRLKDGQTVVTVYADVMSALHAS 147
 QY 62 FKQG-SPEKKLRVREANILMDDIGAEVTPVWDEVGILLHYRMHTELPTFFSSNFD 120
 DB 148 YDQGSGEKFRLCEVDLVLDVDEIGIQTET---KNEQV---VLA-QIVRRRTASMSVGM 201
 QY 121 YSELHHLAMTRDGEKTKVARIIRV 147
 DB 202 LTVMLNTBAMKTLIG-----RIMDKM 222

RESULT 10
 Y4BW_RHISN STANDARD; PRT; 263 AA.

AC P55380;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Putative insertion sequence ATP-binding protein Y4BW/Y4KI/Y4TA.
 GN Y4BW AND Y4KI AND Y4TA.
 OS Rhizobium sp. (strain NGR234).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Feilberg C.A., Fellay R., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";
 RL Nature 387:394-401(1997).
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING
 PROTEIN FAMILY.
 CC -1- SIMILARITY: STRONG, TO Y4UH, ALSO TO Y4IO/Y4ND/Y4SD AND Y4PL.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL: AE000066; AAB91628.1; -
 EMBL: AE000081; AAB91740.1; -
 DR EMBL; AE000097; AAB91855.1; -
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR002611; IetB_Atpbind.
 DR Pfam; PF01695; IetB_1.
 DR SMART; SM00382; AAA; 1.
 KW Hypothetical protein; Plasmid; Transposable element; ATP-binding.
 FT NP BIND 112 119 ATP (POTENTIAL).
 SQ SEQUENCE 263 AA; 29607 MW; 756836C342DC955B CRC64;

Query Match 12.2%; Score 103.5; DB 1; Length 263;
 Best Local Similarity 21.3%; Pred. No. 0.028;
 Matches 37; Conservative 29; Mismatches 55; Indels 53; Gaps 7;

QY 8 AITNGEVQ--KGLVYPPFGTGSFIIIGAIANQLSKKVRSTIYLPEFIRTLKGGF 62
 DB 95 ALAQGGWLRNHLRLITGHTGKSWLACAFGQ--AARLGHVLYVRVPRMFEELALAR 152
 QY 63 KQSGFEKKLRVREANILMDDIG-----AEE-----VTPW 93
 DB 153 LDGSPFLRLDLRLRVQLLLIDDLGTTLSDQGFHLFEIVEERYQRKSTLITQAVVASW 212

QY 94 ---VRBEVIGPLHYRWHELPFFSSNPDYSELEHIAANTROGEEKTKARII 144
 DB 213 HDLIASTVADALIDRIVHN-----AHRTITRGSMKQKSAPIIL 252

RESULT 11

DNA_BUCAL STANDARD; PRT; 454 AA.

ID DNA_BUCAL

AC P57128;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chromosomal replication initiator protein dnaa.

GN DNA OR BU012

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

OS symbiotic bacterium).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Buchnera.

OX NCBI_TaxID=118099;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tokyo 1998;

RA MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids

RT Buchnera sp. APS."

RL Nature 407:81-86(2000).

-1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION

OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;

IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS

(DNA BOX): 5'-TTATC(C/A)(C/A)-3'. DNA BINDS TO ATP AND TO

ACIDIC PHOSPHOLIPIDS.

-1- SIMILARITY: Belongs to the dnaa family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 12.1%; Score 103; DB 1; Length 454;
 Best Local Similarity 32.9%; Pred. No. 0.058;
 Matches 25; Conservative 15; Mismatches 34; Indels 2; Gaps 1;

QY 12 GEGVKGILYLPFGTGSFLIGAIANOLKSKVRSTIIVPE--PIRTLKGFGDGSFEK 69

DB 149 GNSYNPFLGLGAGLGLKTHLHAIHGNILSYKDIKIFMNSSECFVDMVKALKNNAIEK 208

QY 70 KLRVREANILMLDDI 85

DB 209 FKLYRSVDALLDDI 224

RESULT 12
 DNAC_BUCAL STANDARD; PRT; 246 AA.
 AC P57134;

DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA replication protein dnaa.

GN DNAC OR BU021.

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

OS symbiotic bacterium).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Buchnera.

OX NCBI_TaxID=118099;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tokyo 1998;

RA MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids

RT Buchnera sp. APS."

RL Nature 407:81-86(2000).

-1- FUNCTION: THIS PROTEIN IS REQUIRED FOR CHROMOSOMAL REPLICATION. IT

FORMS, IN CONCERT WITH DNA PROTEIN AND OTHER PREPRIMING PROTEINS

OF THE TEMPLATE DNA RECOGNIZED BY PROTEIN N' (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE DNAC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 12.0%; Score 102.5; DB 1; Length 246;
 Best Local Similarity 24.5%; Pred. No. 0.032;
 Matches 34; Conservative 29; Mismatches 61; Indels 15; Gaps 4;

QY 13 EOYKGLYLPFGTGSFLIGAIANOLKSKVRSTIIVPEFIRTLKGFGDGS---FEK 69

DB 97 ENIASFTFSKPGTGTGNHLSAIGNVLILHGSILLYTADLMSNMKTFSSGNSITTEEN 156

QY 70 KLRVREANILMLDDIGAEVTPWVDEVIGPLHYRWHELPFFSSNPDYSELEHIA 129

DB 157 LKHDSSVDLIMIDELGWTGTSRYEK-VIINQIVDRSSSKSTGMLSLNDHKMSLL- 214

QY 130 MTRDGEKTKAARIIEVK 148

DB 215 ---GE-----RVIDRMR 223

RESULT 13
 SKOL_SCHPO STANDARD; PRT; 432 AA.
 ID SKOL_SCHPO
 AC 009803;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Suppressor protein of bent/beds mutants.
 GN SPAC2G11.06.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Valencik M.L., Pringle J.R.;

Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.

[2]
 RN SEQUENCE FROM N.A.
 RP STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajendram M.A., Lyne M., Lyne R., Stewart A.,
 RA Scourto J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle R.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Money P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicikert G., Aert R., Robben J., Glynnpriz B.,
 RA Weljens I., Vanstreelbe E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Filicz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Beger P., Zimmermann W., Medler H., Mambutt R., Purnelle B.,
 RA Goffeau A., Cadieu B., Dreano S., Gloux S., Lelaure V., Mortier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovskiy G.V., Uesery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of *Schizosaccharomyces pombe*."
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES. VERY HIGH TO
 CC MOUSE SKD1 AND YEAST END13.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L33456; AAA35347.1; -;
 DR EMBL: Z54354; CAA91171.1; -;
 DR PIR: T38572; S62461.
 DR GenDB Spombe: SPAC2G11.06; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003959; AAA_ATPase_cent.
 DR InterPro: IPR003960; AAA_sub.
 DR Pfam: PF00004; AAA; 1.
 DR Pfam: PF04212; MIT; 1.
 DR SMART: SMO0382; AAA; 1.
 DR PROSITE: PS00674; AAA; 1.
 DR ATP-binding.
 KM NP_BIND 169 176 ATP (POTENTIAL).
 FT SQUENCE 432 AA; 48401 MW; 65B6948AB9360DIE CRC64;
 SQ
 Query Match 12.0%; Score 102.5; DB 1; Length 432;
 Best Local Similarity 22.4%; Pred. No. 0.061;
 Matches 44; Conservative 31; Mismatches 54; Indels 67; Gaps 9;
 QY 17 GLVLYGPGTGSFLLGAIANQLKSKVRSSTIYLPFRITLKGSGFKDGSFEKLLH 47
 DB 164 GILLYGPGTGSFLLGAIANQLKSKVRSSTIYLPFRITLKGSGFKDGSFEKLLH 223
 QY 48 IYLYPFRITLKGSGFKDGSFEKLLHRYREANITLMDIGAEV-----TPWRYDEV 98
 DB 224 IYLYPFRITLKGSGFKDGSFEKLLHRYREANITLMDIGAEV-----TPWRYDEV 281
 QY 99 IG-----PLLYHYM-----VHELPTEFSSNPDYSELEHILAMRDESEKTKKA 141
 DB 282 IRRRFRKRIYIPLPNMAHARARMPLELNVGKIPSELTSO-DKKE---LAKMTDYSGSDIS 336
 QY 142 -----RIIRRVKSLST 152

DB 337 IYVRAIMEFVRRIHT 352
 RESULT 14
 ID ISTB_BACST STANDARD; PRT; 251 AA.
 AC Q45619;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Insertion sequence IS5376 putative ATP-binding protein.
 OS *Bacillus stearothermophilus*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 OX NCBI_TaxID=1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C021;
 RX MEDLINE=93181468; PubMed=8382825;
 RA Xu K., He Z.-Q., Mao Y.-M., Shen R.-Q., Sheng Z.-J.,
 RT "On two transposable elements from *Bacillus stearothermophilus*."
 RL Plasmid 29.1-9(1993)
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING
 CC PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X67861; CAA48046.1; -;
 DR PIR: S23889; S23889.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR002611; IS2B_ATPbind.
 DR Pfam: PF01695; ISCB; 1.
 DR SMART: SMO0382; AAA; 1.
 KM Transposable element; ATP-binding.
 FT NP_BIND 105 112 ATP (POTENTIAL).
 SQ SQUENCE 251 AA; 29286 MW; FAFBFAF55B2FB527 CRC64;
 Query Match 11.9%; Score 101; DB 1; Length 251;
 Best Local Similarity 20.8%; Pred. No. 0.045;
 Matches 33; Conservative 35; Mismatches 65; Indels 26; Gaps 4;
 QY 13 EQYKGLVLYGPGTGSFLLGAIANQLKSKVRSSTIYLPFRITLKGSGFKDGSFEKLLH 72
 DB 96 DRKENILFPGPGIGKTHLAIISGMEALRAGYKTYFTAHDLVNLBRADQGLKLEKLR 155
 QY 73 RYREANITLMDIGAEVTPWRYDEVYGLLY-----RMVHELPTEFSSNPDYSELEH 126
 DB 156 VVVKPFLVLLIDMGYKLP-----NSAHYLPQVIAIRRYEHAVIILTSNKSFGEW-- 205
 QY 127 HLAMTDGE---EKTVAARIIRVKSISPPYPLSGENFR 162
 DB 206 -----GEIVDSVLAATMDRLHLHGIIPVLSKESYR 237
 RESULT 15
 ID RVUB_MYCGE STANDARD; PRT; 307 AA.
 AC Q49425;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE holliday junction DNA helicase rvub.
 OS RVUB OR MG359.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]


```

RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Frazer C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
CC -1- FUNCTION: THE RUVA-RUVB COMPLEX IN THE PRESENCE OF ATP RENATURES
CC CRUCIFORM STRUCTURE IN SUPERCOILED DNA WITH PALINDROMIC SEQUENCE,
CC INDICATING THAT IT MAY PROMOTE STRAND EXCHANGE REACTIONS IN
CC HOMOLOGOUS RECOMBINATION. RUVAB IS AN HELICASE THAT MEDIATES THE
CC HOLIDAY JUNCTION MIGRATION BY LOCALIZED DENATURATION AND
CC REANNEALING (BY SIMILARITY).
CC -1- SUBUNIT: FORMS A COMPLEX WITH RUVA.
CC -1- SIMILARITY: BELONGS TO THE RUVB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39717; AAC71584.1; -.
DR PIR; G64239; G64239.
DR TIGR; MG359; -.
DR HAMAP; MF_00016; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004605; RuvB.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfam; TIGR00635; ruvB; 1.
DR DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
KW Complete proteome.
FT NP BIND 45 52 ATP (POTENTIAL).
SQ SEQUENCE 307 AA; 35000 MW; 376E5138279E396C CRC64;

Query Match 11.8%; Score 100; DB 1; Length 307;
Best Local Similarity 26.5%; Pred. No. 0.07;
Matches 31; Conservative 27; Mismatches 39; Indels 20; Gaps 5;

Oy 11 NSEOVGLVLYGPGTGTGKFIIGAIANOLSKKVRSTIYLPFIITLKGKGF--KDGSPF 68
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 NKTQLDHILLYGPGYGVKTLTARLAINELKTK-----LQIQGHLQKPSDFL 81
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Oy 69 KILHRYEANIIMLDDIGAEVTPWVRDEVYIGPLHYRMVHELPTFPSSNFDYSLEL 125
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 82 NALSLIKGQVLFIDEIHA--VAPNVM-ELMYPVMD--VFKIQVLIQKDFNSKIYE 132
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: December 15, 2003, 15:14:18
 Job time : 6.18845 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 / Search time 18.9491 Seconds

(without alignments)
2235.384 Million cell updates/sec

Title: US-09-689-952-16

Perfect score: 851
Sequence: 1 AADDICTAITNGEQVKGILYLV.....ERVKSLSTPYFLSGENFRNN 164

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaeoplasmid:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	851	100.0	306	16	Q99TH8 staphylococ
2	851	100.0	306	16	Q8NM67 staphylococ
3	782	91.9	306	16	Q8CNY2 staphylococ
4	423.5	49.8	307	16	Q92BFP4 staphylococ
5	420.5	49.4	311	16	Q9K863 bacillus ha
6	418.5	49.2	308	16	Q8CXB0 oceanobacti
7	418.5	49.2	307	16	Q8Y6X1 listeria mo
8	351	41.2	293	16	Q8P2J5 streptococ
9	351	41.2	300	16	Q9A1D1 streptococ
10	347	40.8	298	16	Q8CWP7 streptococ
11	347	40.8	299	16	Q8CWP7 streptococ
12	346	40.7	298	16	Q97PC7 streptococ
13	332.5	39.1	300	16	Q8S3T8 streptococ
14	332.5	39.1	300	16	Q8DY72 streptococ
15	330	38.8	293	16	Q9CHH9 lactococcus
16	205.5	24.1	74	2	005653 bacillus ce

17	194	22.8	343	16	Q8EMKO	Q8EMKO mycoplasma
18	166	19.5	209	2	Q48991	Q48991 mycoplasma
19	149	17.5	297	2	Q48991	Q48991 mycoplasma
20	140.5	16.5	282	16	Q97H55	Q97H55 clostridium
21	140	16.5	316	16	Q9PQ14	Q9PQ14 ureaplasma
22	129.5	15.2	235	16	Q67056	Q67056 aquifex aeo
23	125.5	14.7	285	16	Q8E236	Q8E236 leptospira
24	122.5	14.4	259	9	Q9G022	Q9G022 bacterioph
25	118.5	13.9	259	9	Q9B0F8	Q9B0F8 staphylococ
26	116	13.6	261	16	Q932A3	Q932A3 staphylococ
27	113.5	13.3	195	17	Q8PTM8	Q8PTM8 methanosa
28	112.5	13.2	257	9	Q9WBR8	Q9WBR8 staphylococ
29	112.5	13.2	257	17	Q8TRM8	Q8TRM8 methanosa
30	112.5	13.2	257	17	Q8TH29	Q8TH29 methanosa
31	111	13.0	350	16	Q92TY8	Q92TY8 rickettsia
32	110	12.9	262	2	Q93P71	Q93P71 microscilla
33	109	12.8	316	2	P95446	P95446 pseudomonas
34	108.5	12.7	327	16	Q8R6U1	Q8R6U1 thermomane
35	108.5	12.6	259	2	Q93SD5	Q93SD5 escherichia
36	107	12.5	261	2	Q8G101	Q8G101 pseudomonas
37	106.5	12.5	261	2	Q9F856	Q9F856 arabidopsis
38	106.5	12.5	1038	10	Q9F856	Q9F856 escherichia
39	105.5	12.4	248	16	Q8X576	Q8X576 pyrococcus
40	105.5	12.4	796	17	Q8TZV0	Q8TZV0 escherichia
41	105	12.3	246	16	Q8XAD9	Q8XAD9 thermoplas
42	105	12.3	375	17	Q97BR9	Q97BR9 escherichia
43	104.5	12.3	248	16	Q8X4T8	Q8X4T8 escherichia
44	104.5	12.3	257	17	Q8TLB4	Q8TLB4 methanosa
45	104.5	12.3	287	16	Q926A4	Q926A4 listeria in

ALIGNMENTS

RESULT 1	ID	Q99TH8	PRELIMINARY:	PRT:	306 AA.
AC	Q99TH8	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)				
DE	Primosomal protein.				
GN	DNAI OR SAV1684 OR SA1507.				
OS	Staphylococcus aureus (strain Mu50 / ATCC 700639), and				
OS	Staphylococcus aureus (strain N315).				
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.				
OX	NCBI TaxID=158879; 158879;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=Staphylococcus aureus (strain Mu50), and S. aureus (strain N315);				
RA	MEDLINE=21311952; PubMed=11418146;				
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,				
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Itoh J.-Q., Ito T.,				
RA	Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,				
RA	Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,				
RA	Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,				
RA	Kaneshima M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,				
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;				
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus				
RT	aureus";				
RL	Lancet 357:1225-1240(2001).				
DR	EMBL; AP003363; BAB57846.1; -				
DR	EMBL; AP003134; BAB42774.1; -				
KW	Complete proteome.				
SC	SEQUENCE 306 AA; 35635 MW; 7F3440E89643505E CRC64;				
Query Match	100.0%; Score 851; DB 16; Length 306;				
Best Local Similarity	Pred. No. 5.6e-73;				
Matches	164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1 AADDICTAITNGEQVKGILYLVGPFGTGSFTLGAIALNQLKSKKRVSTIIYVPEFRTLKG 60				
Db	143 AADDICTAITNGEQVKGILYLVGPFGTGSFTLGAIALNQLKSKKRVSTIIYVPEFRTLKG 202				

Qy 61 GFKDGSFEKKLHVRREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPD 120
 Db 203 GFKDGSFEKKLHVRREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPD 262

Qy 121 YSELEHHLAMTRDGESEKTKAARIETRVKSLSTPYFLSGENFRNN 164
 Db 263 YSELEHHLAMTRDGESEKTKAARIETRVKSLSTPYFLSGENFRNN 306

RESULT 2

Q8NM67 PRELIMINARY; PRT; 306 AA.
 AC Q8NM67;
 DT 01-OCT-2002 (Tremblrel. 22, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Primosomal protein.
 GN Dnal OR MW1627.
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 ON NCBI_TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A., Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L., Yamamoto K., Hiramatsu K.,
 RA "Genome and virulence determinants of high virulence community-acquired MRSA."
 RT Lancet 359:1819-1827 (2002).
 RL EMBL, AP004827; BAB95492.1; -.
 KM Complete proteome.
 SQ SEQUENCE 306 AA; 35626 MW; COFCA3752D934B4 CRC64;

Query Match 100.0%; Score 851; DB 16; Length 306;
 Best Local Similarity 100.0%; Pred. No. 5, 6e-73;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AADICTAINTNGEYVGLYVGPFGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 60
 Db 143 AADICTAINTNGEYVGLYVGPFGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 202

Qy 61 GFKDGSFEKKLHVRREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPD 120
 Db 203 GFKDGSFEKKLHVRREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPD 262

Qy 121 YSELEHHLAMTRDGESEKTKAARIETRVKSLSTPYFLSGENFRNN 164
 Db 263 YSELEHHLAMTRDGESEKTKAARIETRVKSLSTPYFLSGENFRNN 306

RESULT 3

Q8CNV2 PRELIMINARY; PRT; 306 AA.
 AC Q8CNV2;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Primosomal protein.
 GN SEI358.
 OS Staphylococcus epidermidis.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 ON NCBI_TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 12228;
 RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z., Chen Z., Wen Y.,
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 KM Complete proteome.
 SQ SEQUENCE 306 AA; 35736 MW; EDD9F7B52DBFCA16 CRC64;

Query Match 91.9%; Score 782; DB 16; Length 306;
 Best Local Similarity 89.6%; Pred. No. 2, 1e-66;
 Matches 147; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AADICTAINTNGEYVGLYVGPFGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 60
 Db 143 AADICTAINTNGEYVGLYVGPFGTGSFLGAIANOLSKKVRSTIYLPFIRTLKG 202

Qy 61 GFKDGSFEKKLHVRREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPD 120
 Db 203 GFKDGSFEKKLHVRREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPD 262

Qy 121 YSELEHHLAMTRDGESEKTKAARIETRVKSLSTPYFLSGENFRNN 164
 Db 263 YSELEHHLAMTRDGESEKTKAARIETRVKSLSTPYFLSGENFRNN 306

RESULT 4

Q92BF4 PRELIMINARY; PRT; 307 AA.
 AC Q92BF4;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Primosome component (Helicase loader) Dnal.
 GN Dnal OR LIN1595.
 OS Listeria innocua.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 ON NCBI_TaxID=1642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CLIP 11262 / Serovar 6a;
 RX PubMed=11679669;
 RA Glaeser P., Francau L., Buchrieser C., Ruehli C., Amend A., Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T., Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P., Doman E., Dominguez-Bernal G., Duchaud E., Durant L., Duesenget O., Ertan K.-D., Fein H., Garcia-del Portillo F., Garrido P., Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut U., Jackson D., Jones L.-M., Kaerst U., Krell T., Kuhn M., Kunst F., Kutrappat G., Madueno E., Maltounan A., Mata Vicente J., Ng B., Nedjati H., Nordstiek B., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R., Rammel B., Rose M., Schlueter T., Simoes N., Tierrez A., Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 RT "Comparative genomics of Listeria species."
 RL Science 294:849-852 (2001).
 DR EMBL, AL596169; CAC96826.1; -.
 DR Lister: LIN01595;
 DR InterPro: IPR003593; AAA_ATPase.
 DR SMART: SM00382; AAA; 1.
 KM Helicase; Complete proteome.
 SQ SEQUENCE 307 AA; 35988 MW; A8517663BA8F123F CRC64;

Query Match 49.8%; Score 423.5; DB 16; Length 307;
 Best Local Similarity 52.3%; Pred. No. 3, 3e-32;
 Matches 80; Conservative 32; Mismatches 40; Indels 1; Gaps 1;

Qy 11 NGEYVGLYVGPFGTGSFLGAIANOLSKKVRSTIYLPFIRTLKGFKDGSFEKK 70
 Db 154 NGEYVGLYVGPFGTGSFLGAIANOLSKKVRSTIYLPFIRTLKGFKDGSFEKK 213

Qy 71 LHRVREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPDYSELEHHLAM 130
 Db 214 LHRVREANIIMLDIGAEVTPWVRDEVIGPLHYRMVHELPTFFSSNPDYSELEHHLAM 273

Qy 131 TRDG-EKTKAARIETRVKSLSTPYFLSGENFR 162
 Db 274 AONGTEKLAARIETRVKSLSTPYFLSGENFR 306

RESULT 5

Q9K863

```

ID 09K863 PRELIMINARY; PRT; 311 AA.
AC 09K863;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Primosome component (Helicase loader)
GN Dnal OR BH3144.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_Taxid=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maeno N.,
  Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
  Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
  halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AP001517; BAB06863.1; -.
DR InterPro: IPR003593; AAA_ATPase.
DR SMART: SM00382; AAA; 1.
KW Helicase; Complete proteome.
SQ SEQUENCE 311 AA; 35822 MW; 6AFB3073431363BB CRC64;

Query Match 49.4%; Score 420.5; DB 16; Length 311.
Best Local Similarity 50.0%; Pred. No. 6, 5e-32;
Matches 77; Conservative 33; Mismatches 43; Indels 1; Gaps 1;

QY 12 GQVQKLYVYGPFGTGSFLIGAIANOLKSKRVSTIYLPFRITLKGFGKDSFEKKL 71
DB 156 GEDQKLYVYKRGVGTPLMGAINELKDRGIDSTIVYVDFRILKOSIGGTGQGL 215
QY 72 HVRREANIMLDDIGAEVTPWVRDEVIGPLHRYMWHLPFPSSNPFYSELEHLLAMT 131
DB 216 DFKNAQVIFPDIDIGAEVTPWVRDEVIGPLHRYMWHLPFPSSNPFYSELEHLLAMT 275
QY 132 -RDGEKTKAARIERYKSLSTPYPLSGENFRNN 164
DB 276 DKSGTELLAKRMYERIRHYTVSVWVGQGNRRRH 309

RESULT 6
Q8CX80 PRELIMINARY; PRT; 308 AA.
AC 08CX80;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Primosome component (Helicase loader).
GN Dnal OR OB2156.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacilliales; Oceanobacillus.
OX NCBI_Taxid=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=1223376;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
  Ridge and its unexpected adaptive capabilities to extreme
  environments.";
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL: AP004600; BAC14112.1; -.
KW Helicase; Complete proteome.
SQ SEQUENCE 308 AA; 35729 MW; 8D4C2F70E480A833 CRC64;

Query Match 49.4%; Score 420; DB 16; Length 308;
Best Local Similarity 53.4%; Pred. No. 7, 2e-32;
Matches 79; Conservative 30; Mismatches 39; Indels 0; Gaps 0;

QY 16 KGLYVGFPGTGSFLIGAIANOLKSKRVSTIYLPFRITLKGFGKDSFEKKLHRYV 75

```

```

DB 161 KGLYVGFPGTGSFLIGAIANOLKSKRVSTIYLPFRITLKGFGKDSFEKKVDFPK 220
QY 76 EANIMLDDIGAEVTPWVRDEVIGPLHRYMWHLPFPSSNPFYSELEHLLAMTRODE 135
DB 221 KADIMLDDMGAEKMSAMFRDEVLSVLYGRMEGLPVRTSNYDLDLOELSTRNGV 280
QY 136 EKTAKARIERYKSLSTPYPLSGENFRNN 163
DB 281 EGVKGRILLERIKQVTTDVKLSGPNRRS 308

RESULT 7
Q8Y6X1 PRELIMINARY; PRT; 307 AA.
AC 08Y6X1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Primosome component (Helicase loader) Dnal.
GN Dnal OR LM01560.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX NCBI_Taxid=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
  Baquero F., Berche P., Bloesch H., Brandt P., Chakraborty T.,
  Charrif A., Chetoui A., Couve E., de Daruvar A., Deloux P.,
  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussauguet O.,
  Entian K.-D., Feibi H., Garcia-del Portillo F., Garrido P.,
  Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
  Jones R.-M., Kaerst U., Kretz J., Kuhn M., Kunet F., Kurapkhat G.,
  Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
  Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
  Remmel B., Rose M., Schlueter T., Simoes N., Tietz A.,
  Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomes of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL: AL591979; CAC99638.1; -.
DR Listlist; LM001560; -.
DR InterPro: IPR003593; AAA_ATPase.
DR SMART: SM00382; AAA; 1.
KW Helicase; Complete proteome.
SQ SEQUENCE 307 AA; 35970 MW; CF2FD9350CEB0F48 CRC64;

Query Match 49.2%; Score 418.5; DB 16; Length 307;
Best Local Similarity 51.6%; Pred. No. 9, 9e-32;
Matches 79; Conservative 33; Mismatches 40; Indels 1; Gaps 1;

QY 11 NGEVQKLYVGFPGTGSFLIGAIANOLKSKRVSTIYLPFRITLKGFGKDSFEKK 70
DB 154 SGEVYVGLFTHGSGFGKSYLLGALAKELANGISTVLYLPFRMREKVSISDNTVGRK 213
QY 71 LHRVREANIMLDDIGAEVTPWVRDEVIGPLHRYMWHLPFPSSNPFYSELEHLLAM 130
DB 214 IOFAKETVLMDDIGAEVTPWVRDEVIGPLHRYMWHLPFPSSNPFYSELEHLLAM 273
QY 131 TRDG-EKTAKARIERYKSLSTPYPLSGENFR 162
DB 274 AQNGTEKAKARIERYKSLSTPYPLSGENFR 306

RESULT 8
Q8P2J5 PRELIMINARY; PRT; 293 AA.
AC 08P2J5;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Putative primosome component (helicase loader).

```

GN SPY18.0332.
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN NCBI_TaxID=186103;
 [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=MGAS8232 / Serotype M18;
 RX MEDLINE=21927593; PubMed=11917108;
 RA Smoot J.C., Bardian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Studevant D.E., Ricklets S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Karpur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 group A Streptococcus strains associated with acute rheumatic fever
 outbreaks";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 DR EMBL; AB009978; AL97087.1; -
 KM Helicase; Complete proteome; Hypothetical protein.
 SQ SEQUENCE 293 AA; 33357 MW; 0887AD94BDB59986 CRC64;
 Query Match. 41.2%; Score 351; DB 16; Length 293;
 Best Local Similarity 43.1%; Pred. No. 2.6e-25;
 Matches 69; Conservative 36; Mismatches 53; Indels 2; Gaps 2;
 QY 4 DICTAINGEYQVGLYGPFGTGSFLIGAINOLSKK-VSTIYLPFRITLKGCF 62
 DB 135 DPEYQYSAEQ-KGLYVGMGIGKSYLLAMAHLESEKKGVSITLLHPSPFAIDVKNAI 193
 QY 63 KDSGFEKKLRVRANILMLDDIGAEVTPWVRDEVIQPLHYRMVHELPTFSSNPDYS 122
 DB 194 SNGSVKEIDAVKXVPLILDDIGAEQATSWVRREVQLVQLQRMLEELPTFTSNVSPA 253
 QY 123 ELEHNLAMTRDGEKTKAARIIRVKSLSPTYPLSGENFR 162
 DB 254 DLERKVAITIKGSDETWQAKRWERVRYTLAREFHLEGANRR 293
 RESULT 9
 ID Q9A1D1 PRELIMINARY; PRT; 300 AA.
 AC Q9A1D1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Putative primosome component (Helicase loader).
 GN DNA1 OR SPY0340 OR SPY03.0248.
 OS Streptococcus pyogenes, and
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=1314, 198466;
 RN NCBI_TaxID=1314, 198466;
 [1]
 RC SEQUENCE FROM N.A.
 RC SPECIES=S.pyogenes; STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Perrecci J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Szate S., Suvorov A.N., Kenon S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 [2]
 RC SEQUENCE FROM N.A.
 RC SPECIES=S.pyogenes; STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206;
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert F.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 emergence.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 DR EMBL; AB006498; AA33392.1; -
 DR EMBL; AB014141; AA78855.1; -
 KM Hypothetical protein; Helicase; Complete proteome.
 SQ SEQUENCE 300 AA; 34146 MW; 489B27A789338C5A CRC64;
 Query Match. 41.2%; Score 351; DB 16; Length 300;
 Best Local Similarity 43.1%; Pred. No. 2.6e-25;
 Matches 69; Conservative 36; Mismatches 53; Indels 2; Gaps 2;
 QY 4 DICTAINGEYQVGLYGPFGTGSFLIGAINOLSKK-VSTIYLPFRITLKGCF 62
 DB 142 DPEYQYSAEQ-KGLYVGMGIGKSYLLAMAHLESEKKGVSITLLHPSPFAIDVKNAI 200
 QY 63 KDSGFEKKLRVRANILMLDDIGAEVTPWVRDEVIQPLHYRMVHELPTFSSNPDYS 122
 DB 201 SNGSVKEIDAVKXVPLILDDIGAEQATSWVRREVQLVQLQRMLEELPTFTSNVSPA 260
 QY 123 ELEHNLAMTRDGEKTKAARIIRVKSLSPTYPLSGENFR 162
 DB 261 DLERKVAITIKGSDETWQAKRWERVRYTLAREFHLEGANRR 300
 RESULT 10
 ID Q8CWP7 PRELIMINARY; PRT; 298 AA.
 AC Q8CWP7;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Primosome component (Helicase loader).
 GN DNA1 OR SPY155.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OC NCBI_TaxID=171101;
 RN NCBI_TaxID=171101;
 [1]
 RC SEQUENCE FROM N.A.
 RC MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Aldorn W.B. Jr., Arnold J., Blaszcak L.C., Burgett S.,
 RA DeHoff B.S., Batrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA Leblanc D.J., Lee L.N., Lefkowitz S.J., Lu J., Matsushima P.,
 RA McAben S.M., McManney M., Mcleaster K., Mundy C.W., Nicas T.I.,
 RA Norris P.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,
 RA Sun P.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao G.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Kostek P.R. Jr., Skatrud P.L.,
 RA "Genome of the bacterium Streptococcus pneumoniae strain R6";
 RT J. Bacteriol. 183:5709-5717(2001).
 RL EMBL; AB008523; AAL00359.1; -
 KM Helicase; Complete proteome.
 KW SEQUENCE 298 AA; 34574 MW; 3EA8DBBD1DEBDF22 CRC64;
 Query Match. 40.8%; Score 347; DB 16; Length 298;
 Best Local Similarity 44.0%; Pred. No. 6.3e-25;
 Matches 66; Conservative 36; Mismatches 46; Indels 2; Gaps 2;
 QY 15 VGLVLYGPGFGTGSFLIGAINOLSKK-VSTIYLPFRITLKGCFKDSGFEKKLR 73
 DB 149 LKGLVLYGPGFGVGFMAALAHDLSEKKGVSITLLHPSPFIYDVAISDGNVTLVD8 208
 QY 74 VREANILMLDDIGAEVTPWVRDEVIQPLHYRMVHELPTFSSNPDYSLEHNLAMTRD 133
 DB 209 IKLSEVTLILDDIGAEQATSWVRREVQLVQLQRMLEELPTFTSNVSPA 268
 QY 134 GEEKT-KAARIIRVKSLSPTYPLSGENFR 162
 DB 269 GNDETWARVVRIRIYLAETRLSGVNR 298
 RESULT 11
 ID Q8CWP7

ID 08CWM7 PRELIMINARY; PRT; 299 AA.
 AC 08CWM7;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative DNA replication protein, primosome component (Helicase loader)
 DE DNAI OR SMC.1921.
 GN Streptococcus mutans.
 OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1309;
 OX NCBI_TaxID=1309;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=22295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
 RA Li S., Zhu H., Najjar P., Lai H., White J., Roe B.A., Perletti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental
 pathogen."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 DR EMBL; AE015016; AAN59532.1;
 KW Helicase; Complete proteome.
 SQ SEQUENCE 299 AA; 34410 MW; 98CAAB56F552ABD1 CRC64;
 Query Match 40.8%; Score 347; DB 16; Length 299;
 Best Local Similarity 42.5%; Pred. No. 6.3e-25;
 Matches 68; Conservative 37; Mismatches 53; Indels 2; Gaps 2;
 QY 4 DICTAITNEOVKGLVGPFGKSFILGAINOLSKK-VRSITIYLPFRITLKGFGF 62
 DB 141 DVEBOYRPNBQ-KGLVYIGDMGKSYLMAAMHLSBQGAATILHFPSTIDVKNAI 199
 QY 63 KDSEFEKILRVREANIIMLDIGAEVTPWVDEVIQPLHYRMVHLPFPSSNPDYS 122
 DB 200 NGTVKKEIDAVKTDILIDDIGAGOSTSWIRDEVLIQVLYRMLELPFPFTSYSPK 259
 QY 123 ELHHILMTDGEKTKAARIIRVKSISTPYFLSGENR 162
 DB 260 DLEAKLANIKGSDETQAKRVMERIRYLAKEIHLGENRR 299
 RESULT 12
 Q97PC7 PRELIMINARY; PRT; 298 AA.
 AC Q97PC7;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Primosomal protein DnaI.
 GN SP1711.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1313;
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Raddue D.,
 RA Holtzapfel E., Khouri H., Wolf A.M., Ueberback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickenson T., Hickey E.K.,
 RA Holt I.B., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 pneumoniae."
 RL Science 293:498-506(2001).
 DR EMBL; AE007464; AAK75789.1;
 TIGR, SP1711;

KW Complete proteome.
 SQ SEQUENCE 298 AA; 34259 MW; CA5EBCAP9C30059 CRC64;
 Query Match 40.7%; Score 346; DB 16; Length 298;
 Best Local Similarity 44.3%; Pred. No. 7.8e-25;
 Matches 66; Conservative 35; Mismatches 46; Indels 2; Gaps 2;
 QY 16 KGLVYGPFGKSFILGAINOLSKK-VRSITIYLPFRITLKGFGFGSEKKLHRV 74
 DB 150 KGLVYGPFGKSFVVAALADLSEKRGVSTLHIFSDIVDZNAISDGVKTLVDEI 209
 QY 75 REANILMLDDIGAEVTPWVDEVIQPLHYRMVHLPFPSSNPDYSLEHILMTDGG 134
 DB 210 KLSVLIIDDIGAEOSTVWVDEILQVLIQYRMQENLPFPFTSNFDELHGFPAKVKHG 269
 QY 135 EERT-KAARIIRVKSISTPYFLSGENR 162
 DB 270 NDETWEARVMERIRYLAETRLGCVNR 298
 RESULT 13
 Q9E3T8 PRELIMINARY; PRT; 300 AA.
 AC Q9E3T8;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN DNAI OR GBS1668.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=216495;
 OX NCBI_TaxID=216495;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEM316 / Serotype III;
 RX MEDLINE=2242508; PubMed=12354221;
 RA Glaeser P., Ruenick C., Buchrieser C., Chevalier F., Frangoul L.,
 RA Maeder T., Zouine M., Couve E., Lallouf L., Poyart C., Tieu-Cuot P.,
 RA Kunat F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
 RT invasive neonatal disease."
 RL Mol. Microbiol. 45:1499-1513 (2002).
 DR EMBL; AL766852; CAD47327.1;
 DR Sagalier; GDS1668;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 300 AA; 34679 MW; 9C0B4898661661E9 CRC64;
 Query Match 39.1%; Score 332.5; DB 16; Length 300;
 Best Local Similarity 41.9%; Pred. No. 1.5e-23;
 Matches 62; Conservative 36; Mismatches 49; Indels 1; Gaps 1;
 QY 16 KGLVYGPFGKSFILGAINOLSKK-VRSITIYLPFRITLKGFGFGSEKKLHRV 74
 DB 153 KGLVYIGDMGKSYLMAAMHLSBQGAATILHFPSTIDVKNAI 199
 QY 75 REANILMLDDIGAEVTPWVDEVIQPLHYRMVHLPFPSSNPDYSLEHILMTDGG 134
 DB 213 KSVPLIIDDIGAEOSTVWVDEILQVLIQYRMLELPFPFTSNFDELHGFPAKVKHG 272
 QY 135 EERT-KAARIIRVKSISTPYFLSGENR 162
 DB 273 DETWQAKRVMERIRYLAETRLGCVNR 300
 RESULT 14
 Q9DY72 PRELIMINARY; PRT; 300 AA.
 AC Q9DY72;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Primosomal protein DnaI.

GN DNAI OR SAG1621.
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=216466;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=22222988; PubMed=12200547;
 RA Tettein H., Maignan V., Cieslewicz M.J., Eissen J.A., Peterson S.,
 RA Wessels M.R., Paulsen I.T., Nelson K.E., Margalit I., Read T.D.,
 RA Madoff L.C., Wolf A.M., Beaman M.J., Brinkac U.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radu D., Fedorova N.B., Scanlan D., Khouli H., Mulligan S.,
 RA Carty H.A., Cline R.T., Van Aken S.B., Gill J., Scarselli M.,
 RA Jacobini E.T., Bretton C., Gall G., Mariani M., Vegni F., Malone D.,
 RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V Streptococcus agalactiae";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL: AE014265; AAN00485.1; -.
 DR TIGR: SAG1621; -.
 KW Complete proteome.
 SQ SEQUENCE 300 AA; 34663 MW; 9BED36F86B36B3E9 CRC64;

Query Match 39.1%; Score 332.5; DB 16; Length 300;
 Best Local Similarity 41.9%; Pred. No. 1.5e-23;
 Matches 62; Conservative 36; Mismatches 49; Indels 1; Gaps 1;

QY 16 KGLYVGPFGTGSFLIGAIANOLKSKK-VRSITIVLPEFIRTLKGGFKDGSFEKGLHRY 74
 Db 153 KGLYVGDWGVGSKYLMAAARELSERKGVSTLLHFPSPALDVKNALSSGTVDKIDAV 212
 QY 75 REANILMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFSSNPDYSELEHLAMTRDG 134
 Db 213 KSVPIILLDDIGABQATSWVRDELQVILQHRMLELPTFTSNYSFNDLERKMANIKGS 272
 QY 135 EEKTKARITERVKSISTPYFLSGENFR 162
 Db 273 DETWQAKRVERRYVLAIEFHLGPNRR 300

RESULT 15

Q9CHH9 PRELIMINARY; PRT; 293 AA.
 ID 09CHH9;
 AC 09CHH9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Primosomal protein DnaI.
 GN DNAI OR L10752.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallion O., Malarme K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403.";
 RT Genome Res. 11:731-753(2001).
 RL EMBL: AE006308; AAK04850.1; -.
 DR InterPro: IPR003593; AAA_ATPase.
 DR SMART: SM00382; AAA; 1.
 KW Complete proteome.
 SQ SEQUENCE 293 AA; 33904 MW; D775C530A69445B3 CRC64;

Query Match 38.8%; Score 330; DB 16; Length 293;
 Best Local Similarity 43.4%; Pred. No. 2.6e-23;
 Matches 69; Conservative 30; Mismatches 50; Indels 10; Gaps 2;

QY 9 ITNGEQVKGILYVGPFGTGSFLIGAIANOLKSKKVRSTIIVLPEFIRTLKGGFKDGSFE 68
 Db 142 ITNFPKKKGLYVGDGVGKSPMLAMANEIAKKGISTILLHPTPI-----SLLDPD 194
 QY 69 KK---LHRYREANILMLDDIGAEVTPWVRDEVIGPLHYRWVHELPTFSSNPDYSELE 125
 Db 195 NAKVWVNEIKASQVLYLDDIGAEONNNAWVRDSILQVILQHRMQENLPTFTSNLRLMELE 254
 QY 126 HHLAMTRDDEKTKARITERVKSISTPYFLSGENFRN 164
 Db 255 QHLAETKRADEIWPARKVMERVRYLAIEKRLBETNRHD 293

Search completed: December 15, 2003, 15:15:57
 Job time : 19.9491 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 36.4512 Seconds
(without alignments)
1088.625 Million cell updates/sec

Title: US-09-689-952-18

Perfect score: 1328
Sequence: 1 YXDQKHVGHKFPADCNFV.....ERKSLSTPYFLSGENFRNN 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_19Jun03:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1328	100.0	250	22	AA1980.DAT
2	1328	100.0	313	22	AA1981.DAT
3	1189	89.5	299	22	AA1982.DAT
4	1189	89.5	299	22	AA1983.DAT
5	1189	89.5	307	23	AA1984.DAT
6	851	64.1	164	22	AA1985.DAT
7	472.5	35.6	307	23	AA1986.DAT
8	453.5	34.1	155	22	AA1987.DAT
9	397.5	29.9	300	23	AA1988.DAT

10	393	29.6	92	23	ABP09981
11	365	27.5	300	23	ABP27828
12	361	27.2	298	24	ABU02187
13	342	25.8	293	23	ABP54065
14	112	8.4	291	23	ABP53757
15	109.5	8.2	294	23	ABP54759
16	107	8.1	331	23	ABP65341
17	107	8.1	331	23	ABP65341
18	107	8.1	331	23	ABP65341
19	106.5	8.0	630	23	ABP65547
20	105	7.9	248	22	ABP48390
21	105	7.9	798	22	ABP98965
22	104.5	7.9	386	22	ABP96715
23	103	7.8	304	24	ABU01560
24	102.5	7.7	304	23	ABP40308
25	100.5	7.6	784	22	ABP30180
26	100.5	7.6	996	21	ABP91788
27	99.5	7.5	399	21	ABP26569
28	99.5	7.5	399	21	ABP46306
29	99.5	7.5	400	21	ABP43447
30	99	7.5	435	22	ABP25176
31	99	7.5	453	23	ABP30561
32	99	7.5	484	23	ABP27967
33	99	7.5	855	22	ABP96545
34	99	7.5	1377	23	ABP7532
35	99	7.5	2088	22	AAW80124
36	99	7.5	2166	22	AAW79140
37	99	7.5	2356	22	ABG09072
38	98.5	7.4	281	22	ABG05583
39	98.5	7.4	281	22	ABG18846
40	98.5	7.4	453	24	ABU02771
41	98	7.4	467	22	AAU34802
42	98	7.4	479	22	ABG18303
43	97	7.3	340	21	AAV52000
44	97	7.3	340	21	AAV51629
45	96.5	7.3	441	19	AAW80725

ALIGNMENTS

RESULT 1
AAB47316
AAB47316 standard; Protein; 250 AA.
AC AAB47316;
XX
XX
XX 29-AUG-2001 (first entry)
DT
XX
XX Amino acids 64-313 of S. aureus DnaI.
DE
XX
XX DnaI, S. aureus; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
XX screening assay.
XX
XX Staphylococcus aureus.
OS
XX
XX WO200146383-A2.
XX
XX 28-JUN-2001.
XX
XX 21-DEC-2000; 2000MO-US35180.
XX
XX 22-DEC-1999; 99US-0470512.
XX
XX 12-OCT-2000; 2000US-0689952.
XX
XX (PHAG-) PHAGETECH INC.
XX (WILL) WILLIAMS K M.
XX
XX Pelletier J, Gros P, Dubow M;
XX
XX WPI; 2001-418052/44.
XX
XX Novel DnaI polypeptides useful for treating and diagnosing microbial.

PT preferably bacterial, diseases such as those caused by Staphylococcus
 PT aureus -
 XX
 PS Disclosure; Fig 15; 107pp; English.

CC This sequence shows a DnaI polypeptide derived from S. aureus. S.
 CC aureus DnaI interacts with a growth-inhibitory bacteriophage 77 ORF
 CC 104 gene product, to form the basis of a screening assay. DnaI
 CC polypeptides and polynucleotides are useful for treating microbial,
 CC preferably bacterial, especially Staphylococcal, infections. DnaI
 CC polypeptides and polynucleotides are useful for biological, diagnostic,
 CC prophylactic, clinical and therapeutic use, and as components in
 CC databases useful for search analyses as well as in sequence analysis
 CC algorithms.

CC Sequence 250 AA;

Query Match 100.0%; Score 1328; DB 22; Length 250;
 Best Local Similarity 100.0%; Pred. No. 3.6e-129;

Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKQOQHVDGKRPADCPNFKGHPVPELYVNNRIKIRYLQCPCKIKYDEERFPAELITSH 60
 DB 1 YKQOQHVDGKRPADCPNFKGHPVPELYVNNRIKIRYLQCPCKIKYDEERFPAELITSH 60
 QY 61 HMOBDTLNALKDQIYNHNRDLVMAAADICTAINTGEOVKLYLGPFGTGSFTLGA 120
 DB 61 HMOBDTLNALKDQIYNHNRDLVMAAADICTAINTGEOVKLYLGPFGTGSFTLGA 120
 QY 121 IANOLSKKVRSTIILPEFIRTLKGGFKDGSPEKCLHVRBANILMLDDIGAEVTPWV 180
 DB 121 IANOLSKKVRSTIILPEFIRTLKGGFKDGSPEKCLHVRBANILMLDDIGAEVTPWV 180
 QY 181 RDEVIGPLHYRWVHELPTFFSSNFDYSELEHLLAMTRDGEKTKARIIERYKSLSTPY 240
 DB 181 RDEVIGPLHYRWVHELPTFFSSNFDYSELEHLLAMTRDGEKTKARIIERYKSLSTPY 240
 QY 241 FLSGENFRNN 250
 DB 241 FLSGENFRNN 250

RESULT 2

AAB47317
 ID AAB47317 standard; Protein; 313 AA.

AC AAB47317;

DT 29-AUG-2001 (first entry)

XX S. aureus DnaI.

KW DnaI, S. aureus; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
 screening assay.

OS Staphylococcus aureus.

XX MO200146383-A2.

XX 28-JUN-2001.

PF 21-DEC-2000; 2000MO-US35180.

XX 22-DEC-1999; 99US-0470512.

PR 12-OCT-2000; 2000US-0689952.

XX (PHAG-) PHAGETECH INC.

PA (WILL/) WILLIAMS K M.

XX Pelletier J, Gros P, Dubow M;

XX MPI; 2001-418052/44.

DR N-PSDB; AAC86105.

XX Novel DnaI polypeptides useful for treating and diagnosing microbial,
 PT preferably bacterial, diseases such as those caused by Staphylococcus
 PT aureus -
 XX
 PS Disclosure; Fig 1; 107pp; English.

CC This sequence shows DnaI derived from S. aureus. S. aureus DnaI
 CC interacts with a growth-inhibitory bacteriophage 77 ORF 104
 CC gene product, to form the basis of a screening assay. DnaI
 CC polypeptides and polynucleotides are useful for treating microbial,
 CC preferably bacterial, especially Staphylococcal, infections. DnaI
 CC polypeptides and polynucleotides are useful for biological, diagnostic,
 CC prophylactic, clinical and therapeutic use, and as components in
 CC databases useful for search analyses as well as in sequence analysis
 CC algorithms.

CC Sequence 313 AA;

Query Match 100.0%; Score 1328; DB 22; Length 313;
 Best Local Similarity 100.0%; Pred. No. 4.9e-129;

Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKQOQHVDGKRPADCPNFKGHPVPELYVNNRIKIRYLQCPCKIKYDEERFPAELITSH 60
 DB 64 YKQOQHVDGKRPADCPNFKGHPVPELYVNNRIKIRYLQCPCKIKYDEERFPAELITSH 123
 QY 61 HMOBDTLNALKDQIYNHNRDLVMAAADICTAINTGEOVKLYLGPFGTGSFTLGA 120
 DB 124 HMOBDTLNALKDQIYNHNRDLVMAAADICTAINTGEOVKLYLGPFGTGSFTLGA 183
 QY 121 IANOLSKKVRSTIILPEFIRTLKGGFKDGSPEKCLHVRBANILMLDDIGAEVTPWV 180
 DB 184 IANOLSKKVRSTIILPEFIRTLKGGFKDGSPEKCLHVRBANILMLDDIGAEVTPWV 243
 QY 181 RDEVIGPLHYRWVHELPTFFSSNFDYSELEHLLAMTRDGEKTKARIIERYKSLSTPY 240
 DB 244 RDEVIGPLHYRWVHELPTFFSSNFDYSELEHLLAMTRDGEKTKARIIERYKSLSTPY 303
 QY 241 FLSGENFRNN 250
 DB 304 FLSGENFRNN 313

RESULT 3

AAG82049
 ID AAG82049 standard; Protein; 299 AA.

AC AAG82049;

DT 03-SEP-2001 (first entry)

XX S. epidermidis open reading frame protein sequence SEQ ID NO:1192.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;

KW vaccination; endocarditis.

XX Staphylococcus epidermidis.

XX MO200134809-A2.

XX 17-MAY-2001.

PF 09-NOV-2000; 2000MO-US30782.

XX 09-NOV-1999; 99US-0164258.

XX (GLAXO) GLAXO GROUP LTD.

XX Kimmery MJ;

XX MPI; 2001-316495/33.

DR N-PSDB; AAG52899.

XX Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
 PT useful for vaccinating against infections, e.g. endocarditis -
 XX
 PS Claim 18; Page 342; 2189pp; English.

XX AAH5304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AG81454 to AG83120, from *Staphylococcus epidermidis*.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce host cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH5371 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH5091 to
 CC AAH5098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 299 AA;

Query Match 89.5%; Score 1189; DB 22; Length 299;
 Best Local Similarity 87.2%; Pred. No. 1.2e-114;
 Matches 218; Conservative 23; Mismatches 9; Indels 0; Gaps 0;

QY 1 YDQQRHNDGKFPADCPNFKVGHVPELYVDNNRIKIRYLOCCCKIKYDEERREABITSH 60
 DB 50 YDQQRHNDGKFPADCPNFKVGHVPELYVDNNRIKIRYLOCCCKIKYDEERREABITSH 109
 QY 61 HMQRDTLNAKLDIYNNRRDLVMAAADICTALTNGSOVGLYLGPGFGKSFILGA 120
 DB 110 HMQRDTLNAKLDIYNNRRDLVMAAADICTALTNDKRVKGLYLGPGFGKSFILGA 169
 QY 121 IANQKSKVRSSTIYLPFIRITLKGFGDSFEKKLQVRREANILMLDDIGAEVTPMV 180
 DB 170 IANQKSKVRSSTIYLPFIRITLKGFGDSFEKKLQVRREANILMLDDIGAEVTPMV 229
 QY 181 RDEVIGPLIHYRMVHELPTFFSSNFDYSELEHLAMTRDGEKTKAARIIEVKSLSTPY 240
 DB 230 RDEVIGPLIHYRMVHELPTFFSSNFDYSELEHLAMTRDGEKTKAARIIEVKSLSTPY 289
 QY 241 FLGSENFRRN 250
 DB 290 YLTGKNFRRN 299

RESULT 4

AA82337 Standard; Protein; 299 AA.

AA82337;

03-SEP-2001 (first entry)

S. epidermidis open reading frame protein sequence SEQ ID NO:1768.

Staphylococcus epidermidis SRI strain; infection; diagnosis;

Staphylococcus epidermidis.

WO200134809-A2.

17-MAY-2001.

XX

PF 09-NOV-2000; 2000WO-US30782.

XX 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimerly MJ;

XX WPI, 2001-316495/33.

DR N-PSDB; AAH53187.

PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,
 useful for vaccinating against infections, e.g. endocarditis -

PS Claim 18; Page 488; 2189pp; English.

XX AAH5304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AG81454 to AG83120, from *Staphylococcus epidermidis*.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce host cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH5371 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH5091 to
 CC AAH5098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are present in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 299 AA;

Query Match 89.5%; Score 1189; DB 22; Length 299;
 Best Local Similarity 87.2%; Pred. No. 1.2e-114;
 Matches 218; Conservative 23; Mismatches 9; Indels 0; Gaps 0;

QY 1 YDQQRHNDGKFPADCPNFKVGHVPELYVDNNRIKIRYLOCCCKIKYDEERREABITSH 60
 DB 50 YDQQRHNDGKFPADCPNFKVGHVPELYVDNNRIKIRYLOCCCKIKYDEERREABITSH 109
 QY 61 HMQRDTLNAKLDIYNNRRDLVMAAADICTALTNGSOVGLYLGPGFGKSFILGA 120
 DB 110 HMQRDTLNAKLDIYNNRRDLVMAAADICTALTNDKRVKGLYLGPGFGKSFILGA 169
 QY 121 IANQKSKVRSSTIYLPFIRITLKGFGDSFEKKLQVRREANILMLDDIGAEVTPMV 180
 DB 170 IANQKSKVRSSTIYLPFIRITLKGFGDSFEKKLQVRREANILMLDDIGAEVTPMV 229
 QY 181 RDEVIGPLIHYRMVHELPTFFSSNFDYSELEHLAMTRDGEKTKAARIIEVKSLSTPY 240
 DB 230 RDEVIGPLIHYRMVHELPTFFSSNFDYSELEHLAMTRDGEKTKAARIIEVKSLSTPY 289
 QY 241 FLGSENFRRN 250
 DB 290 YLTGKNFRRN 299

RESULT 5

ABP38947 Standard; Protein; 307 AA.

ABP38947;

24-JUL-2002 (first entry)

Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3792.

XX

XX Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
 KW antibacterial; gene therapy.
 XX Staphylococcus epidermidis.
 OS Staphylococcus epidermidis.
 XX US6380370-B1.
 XX 30-APR-2002.
 PD 13-AUG-1998; 98US-0134001.
 XX 14-AUG-1997; 97US-055779P.
 PR 08-NOV-1997; 97US-064964P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA Doucette-Stamm LA, Bush D;
 PI WPI; 2002-381255/41.
 DR N-PSDB; ABN91492.
 XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
 PT polypeptide, useful for diagnosing and treating bacterial infections -
 XX disclosure; SEQ ID 3792; 267pp; English.
 PS ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.
 XX
 XX
 SQ Sequence 307 AA;
 Query Match 89.5%; Score 1189; DB 23; Length 307;
 Best Local Similarity 87.2%; Pred. No. 1.2e-114;
 Matches 218; Conservative 23; Mismatches 9; Indels 0; Gaps 0;
 QY 1 YKQOQHYDGHKADCPNFYKGVHPELYVDNNRIKIRYLOCPCKIKIDEFPFAELITSH 60
 DB 58 YKQOQHYDGHKADCPNFYKGVHPELYVDNNRIKIRYLOCPCKIKIDEFPFAELITSH 117
 QY 61 HMQRDITLNAKLDKIYMMHRRDLVAMAAADICTAITNGEQVGLYLYGPPGTGKSFTLGA 120
 DB 118 HMQRDITLNAKLDKIYMMHRRDLVAMAAADICTAITNGEQVGLYLYGPPGTGKSFTLGA 177
 QY 121 IANQLSKRVSTIILYPERIRTLKGGFKDGFSEKILRVREANIIMLDDIGAEVTPWV 180
 DB 178 IANQLSKRVSTIILYPERIRTLKGGFKDGFSEKILRVREANIIMLDDIGAEVTPWV 237
 QY 181 RDEVIGPLHYRMVHELPTFSSNFYSELEHILAMTRDGEKTKAARIIRVKSISTPY 240
 DB 238 RDEVIGPLHYRMVHELPTFSSNFYSELEHILAMTRDGEKTKAARIIRVKSISTPY 297
 QY 241 FLSGENFRNN 250
 DB 298 YLTGKNFRNN 307

XX Amino acids 150-313 of S. aureus DnaI.
 DE DnaI; S. aureus; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
 KW screening assay.
 XX Staphylococcus aureus.
 OS Staphylococcus aureus.
 XX WO200146383-A2.
 XX 28-JUN-2001.
 PD 21-DEC-2000; 2000WO-US35180.
 XX 22-DEC-1999; 99US-0470512.
 PR 12-OCT-2000; 2000US-0689952.
 XX (PHAG-) PHAGETECH INC.
 PA (WILL/) WILLIAMS K M.
 XX Pelletier J, Gros P, Dubow M;
 PI WPI; 2001-418052/44.
 DR N-PSDB; AAC86104.
 XX Novel DnaI polypeptides useful for treating and diagnosing microbial,
 PT preferably bacterial, diseases such as those caused by Staphylococcus
 PS aureus -
 XX Claim 41; Fig 15; 107pp; English.
 XX This sequence shows a DnaI polypeptide derived from S. aureus. S.
 CC aureus DnaI interacts with a growth-inhibitory bacteriophage 77 ORF
 CC 104 gene product, to form the basis of a screening assay. DnaI
 CC polypeptides and polynucleotides are useful for treating microbial,
 CC preferably bacterial, especially Staphylococcus, infections. DnaI
 CC polypeptides and polynucleotides are useful for biological, diagnostic,
 CC prophylactic, clinical and therapeutic use, and as components in
 CC databases useful for search analyses as well as in sequence analysis
 CC algorithms.
 XX
 XX
 SQ Sequence 164 AA;
 Query Match 64.1%; Score 851; DB 22; Length 164;
 Best Local Similarity 100.0%; Pred. No. 5.6e-80;
 Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 87 AADDICTAITNGEQVGLYLYGPPGTGKSFTLGAIANOLSKKRVSTIILYPERIRTLKG 146
 DB 1 AADDICTAITNGEQVGLYLYGPPGTGKSFTLGAIANOLSKKRVSTIILYPERIRTLKG 60
 QY 147 GFDGSEKILRVREANIIMLDDIGAEVTPWVRDEVIGPLHYRMVHELPTFSSNPD 206
 DB 61 GFDGSEKILRVREANIIMLDDIGAEVTPWVRDEVIGPLHYRMVHELPTFSSNPD 120
 QY 207 YSELEHILAMTRDGEKTKAARIIRVKSISTPYFLSGENFRNN 250
 DB 121 YSELEHILAMTRDGEKTKAARIIRVKSISTPYFLSGENFRNN 164

RESULT 7
 ABB49355
 ID ABB49355 standard; Protein; 307 AA.
 AC ABB49355;
 XX
 DT 05-FEB-2002 (first entry)
 XX
 XX Listeria monocytogenes protein #2059.
 DE Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.
 XX

OS Listeria monocytogenes.
 XX
 PN WO200177335-A2.
 XX
 XX 18-OCT-2001.
 PD
 XX 11-APR-2001; 2001WO-FR01118.
 PF
 XX 11-APR-2000; 2000FR-0004629.
 PR
 XX (INSP) INST PASTEUR.
 PA
 XX Buchrieser C, Frangeul L, Couve E, Rusniok C, Feizi H, Dehoux P,
 PI Dussurget O, Chetoui F, Nedjari H, Glaeser P, Kunst F, Cossette P,
 PI Daniels U, Goebel W, Kreft J, Kuhn M, Ng B, Vazquez-Boland JA,
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
 PI Chakraborty T, Domann R, Hain T, Berche P, Chablit A, Durant L,
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Madueno E, De Pablo B, Wehland J, Kaerst U, Ertian K, Hauf J,
 PI Rose M, Voss H;
 DR WPI; 2002-010914/01.
 XX
 XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides -
 XX
 XX Claim 6; SEQ ID No 2060; 192bp; French.
 PS
 XX
 XX The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 XX Sequence 307 AA;
 SO
 Query Match 35.6%; Score 472.5; DB 23; Length 307;
 Best Local Similarity 42.1%; Pred. No. 2.3e-40;
 Matches 98; Conservative 47; Mismatches 79; Indels 9; Gaps 3;
 QY 22 GAVPELVYNNRKRKIRYLOCPCKIKYDEBFEEALITSHHMDRTNANKLKDYYNNHRDR 81
 DB 77 GAVPKVLNGBERTVYTYPTKEIKEDKRAVRRIRLSLWPKQVYDANLADYTDDESR 136
 QY 82 LDVAMADADDICTAATN-----GEVKGGLYLPFGTGKSFILGAINOLKSKKVRSTIY 136
 DB 137 ---QLALVAYGQFLANNYPKPSGRVKGFLHSGFGKSTLGLAKELAKIGISTIVY 193
 QY 137 LPEFITTLLKGGFQSGFEKKLHVRBANILMLDDIGAEEYTPVVRDVBVGLPLHYMWH 196
 DB 194 LPEFMEVAVQSISDNTVGEKIQAKETEVLMDDIGAESWTAVTRDEVGAILIQFMQES 253
 QY 197 LPTFFSSNPDYSLEHNLAMTRDG-EKTKAARIIRKVSLSPTPYLSEGNFR 248
 DB 254 LPTFFSSNIMDQLENNHMFQNGTEBKAKARRIMERKYLKRVNLEGKRR 306

ID AAG81811 strand; Protein; 155 AA.
 XX
 AC AAG81811;
 XX
 XX 03-SEP-2001 (first entry)
 DT
 XX
 DE S. epidermidis open reading frame protein sequence SEQ ID NO:716.
 XX
 XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX vaccination; endocarditis.
 XX
 XX Staphylococcus epidermidis.
 OS
 XX
 PN WO200134809-A2.
 XX
 PD 17-MAY-2001.
 XX
 XX 09-NOV-2000; 2000WO-US30782.
 PF
 XX 09-NOV-1999; 99US-0164258.
 PR
 XX (GLAX) GLAXO GROUP LTD.
 PA
 XX
 XX Kimmery WJ;
 PI
 XX
 XX WPI; 2001-316495/33.
 DR
 XX N-PSDB; AAH52661.
 DR
 XX
 XX Claim 18; Page 223; 2188bp; English.
 PS
 XX
 XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81811 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH5091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 CC
 XX
 XX Sequence 155 AA;
 SO
 Query Match 34.1%; Score 453.5; DB 22; Length 155;
 Best Local Similarity 80.4%; Pred. No. 8.3e-39;
 Matches 82; Conservative 12; Mismatches 5; Indels 3; Gaps 1;
 QY 1 YDQOQKHVDGHEKACPNVVKHVELYDNNRKRIRYLOCPCKIKYDEBFEEALITSH 60
 DB 50 YDQOQKVDGSHRYDPCPNFVKGHVELYENRIRIRYLPCKCKIKHDERERDSQILSH 109
 QY 61 HMQQDTLANKLKDYYNNHRDLVMAADDDICTAATNGEQVK 102
 DB 110 HMQQDTLANKLKDYYNNRRERDLVMAADDDICTAAT---QIK 148

RESULT 8
 AAG81811

RESULT 9
 ABP27829
 ID ABP27829 strand; Protein; 300 AA.
 XX

AC ABP27829;
 XX 02-JUL-2002 (first entry)
 XX Streptococcus polypeptide SEQ ID NO 4834.
 XX
 XX Streptococcus GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KM group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX Streptococcus pyogenes.
 OS
 XX WO200234771-A2.
 PN 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB04789.
 PF
 XX 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Frazer C;
 PI Tettein H;
 XX
 XX WPI; 2002-352536/38.
 DR N-PSDB; ABN68460.
 XX
 XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 XX Claim 1; Page 3645; 4525pp; English.
 PS
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 XX
 SO Sequence 300 AA;
 Query Match 29.9%; Score 397.5; DB 23; Length 300;
 Best Local Similarity 35.9%; Pred. No. 1.3e-33;
 Matches 84; Conservative 56; Mismatches 91; Indels 3; Gaps 3;
 QY 17 PNTV-KHVELVYDNRKIRIRYIQCCCKTKYDEREALITSHMQRTLNKLDIY 75
 DB 68 PSYIAKQYQPLIMNEGAVSVYLETELVEAQKQAMISRIQLVSPKSYRHTLSDID 127
 QY 76 MNRDRRLDVMAADDTALTNGQVAGLYVPGFGKSPFIILANCKSKK-VASTI 134
 DB 128 VNAASRKEAFSALIDFPEQYPSAQ-KGLVLYGMGKSYLLAAMHSEKKGVSTIL 186
 QY 135 IYVPEFIRTLKGGKQSGFEKTLHRYEANILMDIDGAEVTPWVDEVIGPLIHRMV 194
 DB 187 LHFPSFALDKVAMISNOSVVEEIDAVNVPLIIDDIGAGQANISWVDEVLYQVILQYRML 246

QY 195 HELPTFPSSNFDYSELHHLAMTRDGEKTKARIRRVASTPTPLSGENR 248
 DB 247 BELPTFTSNYSFADLERKWTIKGSDETWQAKRVKRVYVLLAREFLGKANRR 300
 RESULT 10
 ID ABP09981 standard; Protein; 92 AA.
 XX
 XX ABP09981;
 AC
 XX 25-JUN-2002 (first entry)
 DT
 XX Human ORFX protein sequence SEQ ID NO:19944.
 DE
 XX Human, open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KM hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KM degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KM cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KM hypertension; hypothyroidism; cholesterol ester storage disease;
 KM immune deficiency; immune disorder; infectious disease;
 KM autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KM myasthenia gravis.
 XX
 XX Homo sapiens.
 OS
 XX WO200192523-A2.
 PN 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US10836.
 PF
 XX 30-MAY-2000; 2000US-206132P.
 PR 29-AUG-2000; 2000US-228716P.
 PR
 XX (CURA-) CURAGEN CORP.
 PA
 XX Shimkets RA, Leach MD;
 PT WPI; 2002-106308/14.
 DR N-PSDB; ABN25733.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 PS Disclosure; SEQ ID 19944; 1037pp; English.
 XX
 XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABN15762 to ABN27252 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 92 AA; 29.6%; Score 393; DB 23; Length 92;
 XX Query Match Best Local Similarity 81.3%; Pred. No. 7.6e-33;
 Matches 74; Conservative 10; Mismatches 7; Indels 0; Gaps 0;
 QY 100 QVGLVYGFPGKSPFLIGAIANOLSKKVRSTIYLPFRITLKGFGKDSFEKKLAR 159
 Db 2 RVKGLIHGPFCKGKSPFLIKSINOLSKKVRSTIYLPFRITLKGFGKDSFEKKLAR 61
 QY 160 VREANILMLDIGAEEVTPWVRDEVIQPLIH 190
 Db 62 IREANILMLDIGAEEVTPWVRDEVIQPLIH 92
 RESULT 11
 ABP27828
 ID ABP27828 standard; Protein; 300 AA.
 AC ABP27828;
 XX 02-JUL-2002 (first entry)
 DE Streptococcus polyepide SEQ ID NO 4832.
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KM group A streptococcus; Streptococcus pyogenes; antibacterial;
 XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
 OS Streptococcus agalactiae.
 XX MO200234771-A2.
 PD 02-MAY-2002.
 XX 29-OCT-2001; 2001MO-GB04789.
 PF 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 XX 07-MAR-2001; 2001GB-0005640.
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Frazer C;
 PI Tetteijn H;
 DR WPI: 2002-352536/38.
 DR N-PSDB; ABN68459.
 XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 PS Claim 1; Page 3645; 4525pp; English.
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.
 XX Sequence 300 AA; 27.5%; Score 365; DB 23; Length 300;
 XX Query Match Best Local Similarity 32.4%; Pred. No. 3.1e-29;
 Matches 77; Conservative 56; Mismatches 103; Indels 2; Gaps 2;
 QY 12 KPADCNPFYKGVPELYVNNRIKIRYLOCPKIKYKDEERFAELITSHHMDRLTANL 71
 Db 64 KKKDSQYANKGYEPFLVNNRAGVADVSYLETRLEIKQKQALSDRLNVLNPSYNTIM 123
 QY 72 KDLYNNHRDLVMAAADICTAINTGEQVGLYLYGPGTCKSPFLIGAIANOLSKK-V 130
 Db 124 TDFDINNESRMKAMQGLDFVETPSYNN-KGLVLYGDMGVGKSYLMAMAEELSERKV 182
 QY 131 RSTIYLPFRITLKGFGKDSFEKKLARVREANILMLDIGAEEVTPWVRDEVIQPLIH 190
 Db 183 STLLHFPSPFALDVKAISSGTVKDEIDAVKSVPLILDDIGAEOATSWVRDEIILQVILQ 242
 QY 191 YRMVHELPTFPSSNPDYSLEHHLAMTRDGEKTKARLIERVKSISTEYPLSGENR 248
 Db 243 HRMLEELPTFTSNYSFNDLERKMANIKGSDETWQKRVMERVRYLALIEFHLGPNRR 300
 RESULT 12
 ABU02187
 ID ABU02187 standard; Protein; 298 AA.
 XX ABU02187;
 AC ABU02187;
 XX 11-FEB-2003 (first entry)
 DE S. pneumoniae type 4 strain protein from coding region #1764.
 XX Bacterial meningitis; pneumonia; sepsis; otitis media;
 KM ear infection; antiinflammatory; antibacterial; immunostimulant;
 XX auditory; respiratory; gene therapy; vaccine.
 OS Streptococcus pneumoniae type 4 strain.
 XX MO200277021-A2.
 PD 03-OCT-2002.
 XX 27-MAR-2002; 2002MO-IB02163.
 PF 27-MAR-2001; 2001GB-0007658.
 PR (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Masignani V, Tetteijn H, Frazer C;
 PI Masignani V, Tetteijn H;
 DR WPI: 2003-040579/03.
 DR N-PSDB; ABX07476.
 XX New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
 PT or ear infection -
 PS Claim 1; SEQ ID NO 3528; 56pp; English.
 CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS35454. Also included are an antibody which binds one of the
 CC proteins, treating a patient by administering the protein, DNA or
 CC antibody (in a composition), a kit comprising first and second primers,
 CC which are the nucleic acid cited above or fragments between nucleotides

Query Match	27.2%	Score 361;	DB 24;	Length 298;
Best Local Similarity	34.7%	Pred. No. 8e-29;		
Matches 83; Conservative	50;	Mismatches 96;	Indels 10;	Gaps 4

RESULT 13
ABB54065
ID ABB54065 standard; Protein; 293 AA

DT	16-MAY-2002	(first entry)
XX		
DE	Lactococcus lactis	protein dnaI.
XX		
Kw	Biosynthesis, biodegradation, lactic bacterium, yogurt, cheese	
XX		
OS	Lactococcus lactis	IL1403.
XX		
PN	FR280746-A1.	
XX		
PD	12-OCT-2001.	
XX		
XX	11-APR-2000;	2000FR-0004630.
PF		
XX		
XX	11-APR-2000;	2000FR-0004630.
PR		
XX		
XX	(INRG) INRA INST NAT RECH AGRONOMIQUE.	
PA		
PI	Bolotine A, Sorokine A, Renault P, Ehrlich SD,	
XX		
XX	WPI; 2002-043418/06.	

Query Match	25.8%;	Score 342;	DB 23;	Length 293;
Best Local Similarity	33.3%;	Pred. No. 7.3e-27;		
Matches	85;	Conservative	49;	Mismatches 97;
			Indels	24;
			Gaps	6

RESULT 14
ABB53757
ID ABB53757 standard; Protein; 291 AA

DT	16-MAY-2002	(first entry)
XX		
DE	Lactococcus lactis protein p114.	
XX		
KM	Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.	
XX		
OS	Lactococcus lactis IL1403.	
XX		
FN	FR2807446-A1.	
XX		
PD	12-OCT-2001.	
XX		
PF	11-APR-2000; 2000FR-0004630.	
XX		
PR	11-APR-2000; 2000FR-0004630.	
XX		
PA	(INRG) INRA INST NAT RECH AGRONOMIQUE.	
XX		
PI	Bolotline A, Sorokline A, Renault P, Ehrlich SD,	
XX		
DR	WPI; 2002-043418/06.	
XX		

PT New nucleotide sequence useful in the identification of Lactococcus
lactis and related species -

PS Claim 6; SEQ ID No 459; 2504bp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 291 AA;

Query Match 8.4%; Score 112; DB 23; Length 291;
Best Local Similarity 21.8%; Pred. No. 0.005; Mismatches 94; Indels 74; Gaps 11;
Matches 57; Conservative 36; Mismatches 94; Indels 74; Gaps 11;

QY 36 IRYLQPCPKIKYDE---ERFEALITTS-----HMQ-----RDTLN- 68
DB 29 VAY-KHPVSDIDVPLHGLVNDGLQVTSYCECAKEGIYQHQSASAEIKANEDFLNK 87
QY 69 -----AKLDIYMHNRDLVMAADDICTAITNGEQVGLYL 106
DB 88 SKYGRYSILKTQSLVGKSKLMFARFMTFKNGLEQNVLNQORILAREYTOGGRFMTVFV 147
QY 107 YGPFGRKSFILGALINQ---LKSRYASTIYLPFRTLKGGPKDGSFEKGLHR- 159
DB 148 -GGAGGKSHLMALIQEVNENLKDKFSTLFTNISLREIKNSWNYSDTKAEERLTT 206
QY 160 -VEANIILMDDIGAEV---TPWYRDEVIGPLLYHWMVHELPTFFSSNPDYSELEHLL 214
DB 207 LKRTYDLVLIIDGTSTSTSKNSWQ---GVYIYIYARAGNTIITSNLTKEM- 258
QY 215 AMTRDGEKTKARIIRYKS 235
DB 259 ---RSSYDDKIYSRIWEGSKN 276

RESULT 15

ABBS4759
ID ABBS4759 standard; Protein; 294 AA.

AC ABBS4759;

XX 16-MAY-2002 (first entry)

DT Lactococcus lactis protein p1346.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX Lactococcus lactis IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-0004630.

XX 11-APR-2000; 2000FR-0004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bolocrine A, Sorokine A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.
XX New nucleotide sequence useful in the identification of Lactococcus
PT

PT lactic and related species -

PS Claim 6; SEQ ID No 1461; 2504bp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.
CC Note: The sequence data for this patent is based on equivalent patent
CC WO200177334 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 294 AA;

Query Match 8.2%; Score 109.5; DB 23; Length 294;
Best Local Similarity 22.3%; Pred. No. 0.0092; Mismatches 88; Indels 55; Gaps 8;
Matches 50; Conservative 31; Mismatches 88; Indels 55; Gaps 8;

QY 55 ELTSHMQ-----RDTLN-----AKLDIYMHNRDL 83
DB 68 EGIYQHQSASAEIKANEDFLNKSKYKSLKTQSLVGKSKLMFARFMTFKNSLBEQN 127
QY 84 VMAADDICTAITNGEQVGLYLPFGKSFILGALINQ---LKSRYASTIYLPF 139
DB 128 VLNQORILAREYTOGGRFMTVFV-GGAGGKSHLMALIQEVNENLKDKFSTLFTNIS 186
QY 140 FIRTLLKGGPKDGSFEKGLHR---VREANIILMDDIGAEV---TPWYRDEVIGPLLY 191
DB 187 LIREIKNSWNYSDTKAEERLTTMKRYDLVLIIDGTSTSTSKNSWQ---GVYIYI 242
QY 192 RMVHELPTFFSSNPDYSELEHLLAMTRDGEKTKARIIRYKS 235
DB 243 YNAREGNTIITSNLTKEM-----RSSYDDKIYSRIWEGSKN 279

Search completed: December 15, 2003, 15:13:43
Job time : 36.4512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:41, Search time 13.0674 Seconds
(without alignments)
809.474 Million cell updates/sec

Title: US-09-689-952-18

Sequence: 1 YDQOKHYDGHKFPADCPNRY.....EVKSLSTYFLSGENFRN 250

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	ID	Description
1	1328	100.0	313	US-09-470-512A-2	Sequence 2, Appl1
2	1315	99.0	313	US-09-470-512A-10	Sequence 10, Appl1
3	1189	89.5	307	US-09-134-001C-3792	Sequence 3792, Ap
4	537.5	40.5	280	US-09-470-512A-14	Sequence 14, Appl1
5	531.5	40.0	278	US-09-470-512A-15	Sequence 15, Appl1
6	354	26.7	310	US-09-107-532A-5580	Sequence 5580, Ap
7	124	9.3	268	US-09-328-352-5963	Sequence 5963, Ap
8	102.5	7.7	304	US-09-134-001C-5153	Sequence 5153, Ap
9	92.5	7.0	467	US-09-198-452A-443	Sequence 443, App
10	90	6.8	389	US-08-520-170A-13	Sequence 13, Appl1
11	90	6.8	389	US-09-055-699-13	Sequence 13, Appl1
12	90	6.8	389	US-09-273-565-13	Sequence 13, Appl1
13	90	6.8	389	US-09-565-538-13	Sequence 13, Appl1
14	90	6.8	389	US-09-661-468-13	Sequence 13, Appl1
15	90	6.8	389	US-09-976-165-13	Sequence 13, Appl1
16	89.5	6.7	717	US-08-924-629C-5	Sequence 5, Appl1
17	89	6.7	1349	US-08-612-734B-2	Sequence 2, Appl1
18	88	6.6	1334	US-08-996-545-2	Sequence 2, Appl1
19	88	6.6	1334	US-09-328-320-2	Sequence 2, Appl1
20	88	6.6	1438	US-09-209-916-1	Sequence 1, Appl1
21	87.5	6.6	376	US-09-252-991A-24606	Sequence 24606, A
22	87.5	6.6	567	US-09-134-001C-5646	Sequence 5646, Ap
23	87	6.6	872	US-09-198-452A-163	Sequence 163, App
24	86	6.5	468	US-09-134-001C-5156	Sequence 5156, Ap
25	86	6.5	529	US-09-252-991A-18830	Sequence 18830, Ap
26	86	6.5	957	US-09-107-532A-4021	Sequence 4021, Ap
27	85.5	6.4	928	US-08-641-483-4	Sequence 4, Appl1

28	85.5	6.4	928	US-09-382-911-4	Sequence 4, Appl1
29	85.5	6.4	1117	US-08-841-483-6	Sequence 6, Appl1
30	84.5	6.4	1117	US-09-382-911-6	Sequence 6, Appl1
31	84.5	6.4	527	US-09-107-532A-4822	Sequence 4822, Ap
32	84.5	6.4	795	US-09-328-352-6143	Sequence 6143, Ap
33	83.5	6.3	266	US-09-218-363-8	Sequence 8, Appl1
34	83.5	6.3	267	US-09-218-363-10	Sequence 10, Appl1
35	83.5	6.3	290	US-09-218-363-4	Sequence 4, Appl1
36	83.5	6.3	456	US-09-268-364-21	Sequence 21, Appl1
37	83	6.2	308	US-09-149-476-607	Sequence 607, App
38	83	6.2	399	US-09-149-476-474	Sequence 474, App
39	83	6.2	806	US-08-999-774A-6	Sequence 6, Appl1
40	82.5	6.2	427	US-08-996-857-2	Sequence 2, Appl1
41	82.5	6.2	1964	US-08-790-912-3	Sequence 3, Appl1
42	82.5	6.2	2052	US-08-790-912-2	Sequence 2, Appl1
43	82	6.2	227	US-09-291-170A-9	Sequence 9, Appl1
44	82	6.2	227	US-09-724-884-9	Sequence 9, Appl1
45	82	6.2	1471	US-08-683-839B-3	Sequence 3, Appl1

ALIGNMENTS

```

RESULT 1
US-09-470-512A-2
; Sequence 2, Application US/09470512A
; Patent No. 6376652
; GENERAL INFORMATION:
; APPLICANT: Phagotech, Inc.
; TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
; TITLE OF INVENTION: gene and its encoded protein
; FILE REFERENCE: 21715/1000
; CURRENT APPLICATION NUMBER: US/09/470, 512A
; CURRENT FILING DATE: 1999-12-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-470-512A-2

Query Match      100.0%; Score 1328; DB 4; Length 313;
Best Local Similarity 100.0%; Pred. No. 2.5e-141;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 YDQOKHYDGHKFPADCPNRYKIRYLCCKIKYDERFEALITSH 60
DB      64 YDQOKHYDGHKFPADCPNRYKIRYLCCKIKYDERFEALITSH 123
QY      61 HMQDPTLNKLDIYNNHRDLVMAADDICTAITNGEQVGLYVYGFPGKSFILGA 120
DB      124 HMQDPTLNKLDIYNNHRDLVMAADDICTAITNGEQVGLYVYGFPGKSFILGA 183
QY      121 INQOKSKVRBTITTYLPFRITLKGFGDGFEEKLHVRANILMDDIGAEVTPMV 180
DB      184 INQOKSKVRBTITTYLPFRITLKGFGDGFEEKLHVRANILMDDIGAEVTPMV 243
QY      181 RDEVIGPLHYMVEHLPTFFSSNFDYSELHHLAMTRDGEKTKAARIIEVKSLSTY 240
DB      244 RDEVIGPLHYMVEHLPTFFSSNFDYSELHHLAMTRDGEKTKAARIIEVKSLSTY 303
QY      241 FLSGENFRN 250
DB      304 FLSGENFRN 313

RESULT 2
US-09-470-512A-10
; Sequence 10, Application US/09470512A
; Patent No. 6376652
; GENERAL INFORMATION:
; APPLICANT: Phagotech, Inc.

```

;; TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
;; TITLE OF INVENTION: gene and its encoded protein
;; FILE REFERENCE: 21715/1000
;; CURRENT APPLICATION NUMBER: US/09/470,512A
;; CURRENT FILING DATE: 1999-12-12
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 10
;; LENGTH: 313
;; TYPE: PRT
;; ORGANISM: Staphylococcus aureus
US-09-470-512A-10

Query Match 99.0%; Score 1315; DB 4; Length 313;
Best Local Similarity 98.4%; Pred. No. 7,3e-140;
Matches 246; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKQOQKHVDGKADCPNFYKGVHPELYVNNRIKIRYLQCPCKIKYDEERFEAEILITSH 60
DB 64 YKQOQKHVDGKADCPNFYKGVHPELYVNNRIKIRYLQCPCKIKYDEERFEAEILITSH 123
QY 61 HMQORDTLNAKLDIYNNHRLDVAMAADDICTAINTGEOVKGLYLYGPGTGKSPILGA 120
DB 124 HMQORDTLNAKLDIYNNHRLDVAMAADDICTAINTGEOVKGLYLYGPGTGKSPILGA 183
QY 121 IANQLSKKVRSTIIVLPEFIRTLKGGFPGKSGFEKTLHVRREANIIMLDIGAEVTPWV 180
DB 184 IANQLSKKVRSTIIVLPEFIRTLKGGFPGKSGFEKTLHVRREANIIMLDIGAEVTPWV 243
QY 181 RDEVIGPLHYRMVHELPTFFSSNFDYSELEHHLAMTRDGEETKARIIERVKSLSTPY 240
DB 244 RDEVIGPLHYRMVHELPTFFSSNFDYSELEHHLAMTRDGEETKARIIERVKSLSTPY 303
QY 241 FLSGENFRNN 250
DB 304 FLSGENFRNN 313

RESULT 3

US-09-134-001C-3792
; Sequence 3792, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lyta Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3792
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3792

Query Match 89.5%; Score 1189; DB 4; Length 307;
Best Local Similarity 87.2%; Pred. No. 1.1e-125;
Matches 218; Conservative 23; Mismatches 9; Indels 0; Gaps 0;

QY 1 YKQOQKHVDGKADCPNFYKGVHPELYVNNRIKIRYLQCPCKIKYDEERFEAEILITSH 60
DB 58 YKQOQKHVDGKADCPNFYKGVHPELYVNNRIKIRYLQCPCKIKYDEERFEAEILITSH 117
QY 61 HMQORDTLNAKLDIYNNHRLDVAMAADDICTAINTGEOVKGLYLYGPGTGKSPILGA 120
DB 118 HMQORDTLNAKLDIYNNHRLDVAMAADDICTAINTGEOVKGLYLYGPGTGKSPILGA 177
QY 121 IANQLSKKVRSTIIVLPEFIRTLKGGFPGKSGFEKTLHVRREANIIMLDIGAEVTPWV 180

DB 178 IANQLSKKVRSTIIVLPEFIRTLKGGFPGKSGFEKTLHVRREANIIMLDIGAEVTPWV 237
QY 181 RDEVIGPLHYRMVHELPTFFSSNFDYSELEHHLAMTRDGEETKARIIERVKSLSTPY 240
DB 238 RDEVIGPLHYRMVHELPTFFSSNFDYSELEHHLAMTRDGEETKARIIERVKSLSTPY 297
QY 241 FLSGENFRNN 250
DB 298 FLSGENFRNN 307

RESULT 4

US-09-470-512A-14
; Sequence 14, Application US/09470512A
; Patent No. 6376652
; GENERAL INFORMATION:
; APPLICANT: PhageTech, Inc.
; TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
; TITLE OF INVENTION: gene and its encoded protein
; FILE REFERENCE: 21715/1000
; CURRENT APPLICATION NUMBER: US/09/470,512A
; CURRENT FILING DATE: 1999-12-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-470-512A-14

Query Match 40.5%; Score 537.5; DB 4; Length 280;
Best Local Similarity 44.9%; Pred. No. 2.8e-52;
Matches 106; Conservative 48; Mismatches 81; Indels 1; Gaps 1;

QY 15 DCPNFYKGVHPELYVNNRIKIRYLQCPCKIKYDEERFEAEILITSHMQORDTLNAKLDI 74
DB 45 NQNNLEGHVPLLVNVRSIDIEYECVYKRLDQKQOSLMKSWYIQDILGATFGQV 104
QY 75 YNNHRLDVAMAADDICTAINTGEOVKGLYLYGPGTGKSPILGANQLSKKVRSTI 134
DB 105 DISDPSRLMPOHVTDPLKSNVETGKGLYLYGKGVGTPTMLAIANLEKEYSMTI 164
QY 135 IYLPFIRTLKGGFPGKSGFEKTLHVRREANIIMLDIGAEVTPWVRDEVIGPLHYRMV 194
DB 165 VYVEPVRKLSLDQDTIEKLNWKTTPVIMLDIGAESVTSNVRDEVITGLQHRNS 224
QY 195 HELPTFFSSNFDYSELEHHLAMTRDGE-EKTKARIIERVKSLSTPYFLSGENFRN 249
DB 225 QCLPTFFSSNFDYSELEHHLAMTRDGE-EKTKARIIERVKSLSTPYFLSGENFRN 280

RESULT 5

US-09-470-512A-15
; Sequence 15, Application US/09470512A
; Patent No. 6376652
; GENERAL INFORMATION:
; APPLICANT: PhageTech, Inc.
; TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
; TITLE OF INVENTION: gene and its encoded protein
; FILE REFERENCE: 21715/1000
; CURRENT APPLICATION NUMBER: US/09/470,512A
; CURRENT FILING DATE: 1999-12-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-09-470-512A-15

Query Match 40.0%; Score 531.5; DB 4; Length 278;
Best Local Similarity 41.2%; Pred. No. 1.3e-51;

SEQ ID NO 5153
 LENGTH: 304
 TYPE: PRF
 ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-5153

Query Match
 Best Local Similarity 21.4%; Score 102.5; DB 4; Length 304;
 Best Local Similarity 21.4%; Pred. No. 0.0035;
 Matches 33; Conservative 35; Mismatches 59; Indels 27; Gaps 4;

QY 88 ADDICATINNGEBOVKGLYGPPTGKSPFLGAIANOLSKKVRSTIYLP-----EPI 141
 DB 135 AASIAVAEAPAEKAVNPLFIYGVGLGKTHLMAIGHVLSNKNPAKVIYTSSEKFTWEFI 194
 QY 142 RTKAGGFKDGSPEKHLRVREANIIMDDIGAEVTPWVDEVIGPLHVRMHELPTFP 201
 DB 195 KSI-----RDNETAFREKRYKIVLIDTIOPTONKEQTDE-----FP 234
 QY 202 SSNPDYSELEHHLAMTFDGEKTKAARILIRKVS 235
 DB 235 HTNELHNNKQIVISSDRPK-EIAKLEDRLS 267

RESULT 9
 US-09-198-452A-443
 Sequence 443, Application US/09198452A
 Patent No. 6559294
 GENERAL INFORMATION:
 APPLICANT: Griffiths, R.
 TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
 TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
 TITLE OF INVENTION: and treatment of infection
 FILE REFERENCE: 9710-003-999
 CURRENT APPLICATION NUMBER: US/09/198,452A
 CURRENT FILING DATE: 1998-11-24
 NUMBER OF SEQ ID NOS: 6849
 SEQ ID NO 443
 LENGTH: 467
 TYPE: PRF
 ORGANISM: Chlamydia pneumoniae
 US-09-198-452A-443

Query Match
 Best Local Similarity 21.0%; Score 92.5; DB 4; Length 467;
 Best Local Similarity 21.0%; Pred. No. 0.092;
 Matches 38; Conservative 39; Mismatches 73; Indels 31; Gaps 8;

QY 20 VKGVELVYNNRIKIRYLQCPCK-IKDEERFEA-ELITSHR-----MORDTLNA 69
 DB 60 IRLEVENIIFQNYILD-NYGRDLCSPFLDVHGEPALFVVAHKKPSAPVASQKES-NE 117
 QY 70 KLDIYMNHRD---RDLVMAADDICATIN-----GEQVKGILYGPPTG 112
 DB 118 GISVEFEETDFELKLNLSYRPNFLGSPNQFVKSAAVGIACKPGSYNPLFIHGCVGL 177
 QY 113 GKSFITGAIANOLSKS---KKVRSTIYLPFIYGVGLGKTHLMAIGHVLSNKNPAKVIYTSSEKFTWEFI 170
 DB 178 GKTHLHVAHYVREHKKNLRIHCITTEATINDLVHLKSKSYDKMKNFYSIDLLVVD 237

QY 171 I 171
 DB 238 I 238

RESULT 10
 US-08-820-170A-13
 Sequence 13, Application US/08820170A
 Patent No. 5831058
 GENERAL INFORMATION:
 APPLICANT: Tsutomu, FUJIWARA
 APPLICANT: Takeshi, WATANABE
 APPLICANT: Masato, HORIE
 APPLICANT: Toyomasa, KATAGIRI
 TITLE OF INVENTION: HUMAN GENE

NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Sughrue, Mion, Zinn, Macpeak & Seas
 STREET: 2100 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States
 ZIP: 20037-3202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/820,170A
 FILING DATE:
 CLASSIFICATION: 536
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 293-7060
 TELEFAX: (202) 293-7860
 TELEX: 6491103
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 389 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-820-170A-13

Query Match
 Best Local Similarity 35.1%; Score 90; DB 2; Length 389;
 Best Local Similarity 35.1%; Pred. No. 0.13;
 Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 102 KGLIYGPPTGKSPFLGAIANOLSKS---KKVRSTIYLPFIYGVGLGKTHLMAIGHVLSNKNPAKVIYTSSEKFTWEFI 156
 DB 168 KGLIYGPPTGKSPFLGAIANOLSKS---KKVRSTIYLPFIYGVGLGKTHLMAIGHVLSNKNPAKVIYTSSEKFTWEFI 220
 QY 157 LHRVREANIIMDDIGA 173
 DB 221 YARDHOPCLIFMDEIDA 237

RESULT 11
 US-09-055-699-13
 Sequence 13, Application US/09055699
 Patent No. 6005088
 GENERAL INFORMATION:
 APPLICANT: Tsutomu, FUJIWARA
 APPLICANT: Takeshi, WATANABE
 APPLICANT: Masato, HORIE
 APPLICANT: Toyomasa, KATAGIRI
 TITLE OF INVENTION: HUMAN GENE
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Sughrue, Mion, Zinn, Macpeak & Seas
 STREET: 2100 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States
 ZIP: 20037-3202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/055,699
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/820,170
 FILING DATE:
 TELECOMMUNICATION INFORMATION:

RESULT 13
US-09-565-538-13
Sequence 13, Application US/09565538
Patent No. 6333404
GENERAL INFORMATION:
APPLICANT: FUJIMURA, TSUTOMU
APPLICANT: WATANABE, TAKESHI

Query Match 6.8%; Score 90; DB 4; Length 389;
Best Local Similarity 35.1%; Pred. No. 0.13;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3

102 KALYLPGPFGTGSFIIICAIINOLKS--KVRSTII--YLPEFRITLKGGFGKDSFEKK 156
168 KGCILYLPBPFGSKILLARAVASQDCCNPLKRVSSIVDKYIGESARLLIREF-----N 220

```
Qy      157 LHRVREANILMLDDIGA 173
      | : | : | |
Db      221 YARDHQPCITFMEIDA 237
```

RESULT 15

```

US-09-976-165-13
Sequence 13, Application US/0976165
Patent No. 6562947
GENERAL INFORMATION:
APPLICANT: FUJIMURA, TSUTOMU
APPLICANT: WATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/976,165
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/565,538
PRIOR FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 62410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 65163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 369
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-165-13

```

Query Match	6.8%	Score 90	DB 4	Length 389
Best Local Similarity	35.1%	Pred. No. 0.13		
Matches 27	Conservative 13	Mismatches 25	Indels 12	Gaps 3

```

Qy      102 KGLVYVYFGFGTSPFLGIANQLKS---KVRSTI--YLPEFIRTLKGFQDGSFEKK 156
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      168 KGCLLYGPGTGTCTLLARAVASQDLCNPLKVVSSIVDKTIGESALIREMF-----N 220
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      157 LHRVRANILMDIDIGA 173
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      221 YARDHQPCITFMEIDA 237
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: December 15, 2003, 15:17:40
Job time : 14.0674 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:13:51 / Search time 24.4154 Seconds
(without alignments)
1904.368 Million cell updates/sec

Title: US-09-689-952-18
Perfect score: 1328
Sequence: 1 YDQOKHYDGHKFAKDCPNFV.....ERVKSLSTPYFLSGENFRNN 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues
Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	7.9	248	12	US-10-287-274-435
2	100	7.5	463	15	US-10-156-761-11850
3	98	7.4	657	9	US-09-815-242-10395
4	96.5	7.3	442	9	US-09-815-242-13552
5	96	7.2	454	9	US-09-815-242-11167
6	93.5	7.0	796	12	US-10-032-585-7201
7	92.5	7.0	574	10	US-09-764-868-774
8	91.5	6.9	827	12	US-10-032-585-7621
9	90.5	6.8	577	9	US-09-815-242-10193
10	90	6.8	389	10	US-09-976-165-113
11	90	6.8	389	12	US-10-205-219-165
12	90	6.8	389	12	US-10-342-276-113
13	90	6.8	393	15	US-10-128-714-3421
14	90	6.8	445	9	US-09-815-242-5010
15	90	6.8	447	9	US-09-815-242-10908

16	90	6.8	451	15	US-10-128-714-8421	Sequence 8421, App
17	89.5	6.7	717	11	US-09-883-343A-5	Sequence 5, Appl1
18	88	6.6	444	12	US-10-032-585-7385	Sequence 7385, Ap
19	88	6.6	453	9	US-09-815-242-5668	Sequence 5668, Ap
20	88	6.6	453	9	US-09-815-242-12701	Sequence 12701, A
21	88	6.6	1334	10	US-09-758-828-2	Sequence 2, Appl1
22	88	6.6	1438	14	US-10-006-091-1	Sequence 1, Appl1
23	88	6.6	1438	14	US-10-047-257-1	Sequence 1, Appl1
24	88	6.6	1438	14	US-10-225-960-1	Sequence 1, Appl1
25	87.5	6.6	411	12	US-10-032-585-7284	Sequence 7284, Ap
26	87	6.6	682	12	US-10-032-585-7047	Sequence 7047, Ap
27	87	6.6	866	12	US-09-841-260-113	Sequence 113, App
28	87	6.6	866	14	US-10-007-693-113	Sequence 113, App
29	86.5	6.5	606	15	US-10-128-714-8064	Sequence 3064, Ap
30	86.5	6.5	847	15	US-10-128-714-8064	Sequence 8064, Ap
31	86	6.5	439	15	US-10-128-714-3522	Sequence 3522, Ap
32	86	6.5	439	15	US-10-128-714-3522	Sequence 8522, Ap
33	86	6.5	514	9	US-09-815-242-11772	Sequence 11772, A
34	85.5	6.4	298	12	US-10-380-558-17	Sequence 17, Appl
35	85.5	6.4	444	15	US-10-153-668-308	Sequence 308, App
36	85.5	6.4	465	9	US-09-925-302-771	Sequence 771, App
37	85.5	6.4	482	10	US-09-764-868-1149	Sequence 1149, Ap
38	85.5	6.4	489	10	US-09-764-868-776	Sequence 776, App
39	85	6.4	421	15	US-10-128-714-8485	Sequence 8485, Ap
40	85	6.4	465	15	US-10-128-714-3485	Sequence 3485, Ap
41	84.5	6.4	442	11	US-09-828-062-7	Sequence 7, Appl1
42	84.5	6.4	584	10	US-09-764-868-1054	Sequence 1054, Ap
43	84	6.3	401	12	US-10-032-585-7325	Sequence 7325, Ap
44	84	6.3	508	11	US-09-769-787-102	Sequence 102, App
45	84	6.3	561	15	US-10-156-761-13374	Sequence 13374, A

ALIGNMENTS

RESULT 1
US-10-287-274-435
; Sequence 435, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERETO
; FILE REFERENCE: ELIIRA.008DVI
; CURRENT APPLICATION NUMBER: US/10/287, 274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 435
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-435
Query Match 7.9%; Score 105; DB 12; Length 248;
Best Local Similarity 26.7%; Pred. No. 0.004;
Matches 43; Conservative 30; Mismatches 66; Indels 22; Gaps 7;
QY 77 NHRDLVAMA--ADDICTAITNGEQVGLYLYPEFGTGSFIIICAIINOLKSKVRBT 133
DB 80 NCGQYVALSQAISIDELMTGCTN-----FAFSGPGTGKXHLAALGNRLKQGTVI 133
QY 134 IYVLEPIRTLKSGKDG-SPEKKIHYREANIIMLDIGAEVTPWVDEVIPLHYR 192
DB 134 VVTVAQVMSALASTDDGSGEKFLRELCEVDLVDELGIQRET---KNEQV--VLH-Q 187
QY 193 WVEHLPTFPSSNPDYSELEHILAMTRDGEKTKAARIERV 233

Db 188 IVDRTASMSVGLTNVYEAAMKTLGGE-----RINDRM 222

RESULT 2

US-10-156-761-11850

Sequence 11850, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHITKA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 11850

LENGTH: 653

TYPE: PRF

ORGANISM: Streptomyces avermitilis

US-10-156-761-11850

Query Match

Best Local Similarity 39.2%; Score 100; DB 15; Length 653;

Matches 29; Conservative 15; Mismatches 20; Indels 10; Gaps 4;

QY 104 LHYGPGTGKSTILGAINOLKS---KKVR--STIIVPEFIRLKGKGFKGSFEKRL 157

DB 350 LFIYSGGKGTLLHLAGVARSIVYSGTRVYVSEEFNEFINISIRG-KGDSFRK-- 406

QY 158 HRYREANILMLDDI 171

DB 407 -RYREMDILLVDDI 419

RESULT 3

US-09-815-242-10395

Sequence 10395, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyekind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: PaetSeq for Windows Version 4.0

SEQ ID NO 10395

LENGTH: 467

TYPE: PRF

ORGANISM: Escherichia coli

US-09-815-242-10395

Query Match

Best Local Similarity 7.4%; Score 98; DB 9; Length 467;

Matches 19; Conservative 21; Mismatches 28; Indels 2; Gaps 1;

QY 104 LHYGPGTGKSTILGAINOLKSRYSTIIVL--PEFIRLKGKGFKGSFEKRLHVR 161

DB 168 LFIYSGGKGTLLHLAGVARSIVYSGTRVYVSEEFNEFINISIRG-KGDSFRK-- 227

QY 162 EANIILMLDDI 171

DB 228 SVDALLIIDI 237

RESULT 4

US-09-815-242-13552

Sequence 13552, Application US/09815242

Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyekind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: PaetSeq for Windows Version 4.0

SEQ ID NO 13552

LENGTH: 442

TYPE: PRF

ORGANISM: Streptococcus pneumoniae

US-09-815-242-13552

Query Match

Best Local Similarity 7.3%; Score 96.5; DB 9; Length 442;

Matches 42; Conservative 38; Mismatches 78; Indels 49; Gaps 7;

QY 50 ERFPAELITSHM---QRTIAKLK-----DYMNRDRRLDVAM 86

DB 56 EIVDAE-ITPHYITFKPKQDTSSQVEATNLTLYDVSPLVSIPIVSDTGLKEKYTFDNT 114

QY 87 AADD-----ICTAITNGEOV-----KGLVYGPFGTGKSTILGAINOLKSRYSTIIVL 138

Db 115 QODGNWAVASALAVSEDLATITNPLFIYGPGLGKTHLNLAINIGELNKNIPARVKYIP 174
Qy 139 ---EFIRTLKGFGDGSFEKKLHVRREANILMLDDIGAEVTPWVRDEVIGPLHYMVE 196
Db 175 AASFINDPLDHLRLGEMERFKTYRSLDLDLDDIOS-----LSGKXVAT 219
Qy 197 LPTFSSNFDYSELEHNLAMTRDGEERK 223
Db 220 QEEFNTFMAHLDDKKQIVLTSRSPK 246

RESULT 5

US-09-815-242-11167
Sequence 11167, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trewick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11167
LENGTH: 454
TYPE: PRY
ORGANISM: Haemophilus influenzae
US-09-815-242-11167

Query Match 7.2%; Score 96; DB 9; Length 454;
Best Local Similarity 24.0%; Pred. No. 0.095;
Matches 46; Conservative 31; Mismatches 75; Indels 40; Gaps 9;

Qy 51 REAEELISHHQRDTLNKLDIYNHNRDLVDVMAADDICTAINGQVKGILYIGPF 110
Db 114 KRESHLNTH-----LFDNFVEGKSN-QLAAVGQKLAQAGEPSANPFPLYGCT 162
Qy 111 GTGKFFILGAIANOLSKKVRSTIYL--PEFIRTLKGFGDGSFE--KKLHVRREANIL 166
Db 163 GIGKTHLHLAIGNGLIADKPNARVLYIHANNFMQHWKAVRDNKDDQFKFTRSLDA--L 220
Qy 167 MLDI---GAEEVTPWVRDEVIGPLHYMVEHLPTFSSNFDYSELEHNLAMTRDGEERK 223
Db 221 LVDDIDQFAKEKEKTQ-----EEFHIIFNSLF---ETGQIILTSRSPK 261

Qy 224 TKAARLIERVKS 235
Db 262 -EIEKIEERLKS 272

RESULT 6

US-10-032-585-7201
Sequence 7201, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jians
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7201
LENGTH: 796
TYPE: PRY
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (383)..(383)
OTHER INFORMATION: X-any amino acid
US-10-032-585-7201

Query Match 7.0%; Score 93.5; DB 12; Length 796;
Best Local Similarity 26.5%; Pred. No. 0.41;
Matches 40; Conservative 22; Mismatches 44; Indels 45; Gaps 8;

Qy 67 LNAKLDIY---MNR-----DRLDVAMADDICTAINGEQVKGILYIGPFGTGKSP 116
Db 282 LDAEFQDIFRAFNRIILPPELAEKLDY-----KHCKGILLYGPPGTGKTS 327
Qy 117 ILGAIANOLSKKVRSTIYLPEFIRTLKG-----FKGSEKKLH-RVREANILM 167
Db 328 IARKLSKMLNGKEPK--IVNGPEMLSKYVGASEENIRMLFDAAEAYKLKGEDSLHXII 385
Qy 168 LDDI-----GAEEVTPWVRDEVIGPL 169
Db 386 FDELDSVFKQSGSGSDGT-GVGDVAVNQL 415

RESULT 7

US-09-764-868-774
Sequence 774, Application US/09764868
Patent No. US20020168711A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PAM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 774
LENGTH: 574
TYPE: PRY
ORGANISM: Homo sapiens
US-09-764-868-774

Query Match 7.0%; Score 92.5; DB 10; Length 574;
Best Local Similarity 30.3%; Pred. No. 0.33;
Matches 30; Conservative 17; Mismatches 33; Indels 19; Gaps 4;

Qy 102 KELYVYGFPGTGSFIIIGAIANOLSKKVRSTIYLPEFIRTLKGFGDGSFEKKLHVR- 160
Db 127 RGVLLYGPFGTGKTMILAAVANEVGA---YYSVINGPILISKFY-----GETEALRQIF 178
Qy 161 -----REANILMLDDIGA-----EEVTPWVRDEVIGPL 189
Db 179 AEATLRHPSILIFIDELALCPREBGAONEVEKRVVASILL 217

RESULT 8
US-10-032-585-7621
Sequence 7621, Application US/10032585
Publication No. US20030180953A1
GENERAL INFORMATION:
APPLICANT: Terry, Roemer D.
APPLICANT: Bo, Jiang
APPLICANT: Charles, Boone
APPLICANT: Howard, Bussey
TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
FILE REFERENCE: 10182-005-999
CURRENT APPLICATION NUMBER: US/10/032,585
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 8000
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7621
LENGTH: 827
TYPE: PRT
ORGANISM: Candida albicans
US-10-032-585-7621

Query Match
Best Local Similarity 31.8%; Score 91.5; DB 12; Length 827;
Pred. No. 0.71;
Matches 35; Conservative 17; Mismatches 41; Indels 17; Gaps 6;

QY 96 TNGEQVAGLYGPGTGTGKSFILGAIANOLSKKVRSTIYLPFIRTLKGFQDSFEK 155
DB 222 TNGEPRGVLLYGPBGSKTTIANALAGELKVPIN--ISASVSGM-----SGSEK 273

QY 156 KLRHV-REAN-----IMLDDIGAEBVTWVRDEVIQPLHYRMVHELPT 199
DB 274 KLRIFEAKQIAICLIFMDEIDA--ITP-KRDGGAQRMKRIVAQLLT 320

RESULT 9
US-09-815-242-10193
Sequence 10193, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Hasebeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: PASTESEQ for Windows Version 4.0
SEQ ID NO 10193
LENGTH: 577
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10193

Query Match
Best Local Similarity 18.0%; Score 90.5; DB 9; Length 577;
Pred. No. 0.55;
Matches 68; Conservative 37; Mismatches 86; Indels 187; Gaps 14;

QY 1 YXDQQRHYD-GHKPAD-CENFYKGVHVELYVDNNRIKIRYLCQCKIKYDEERFEABLIT 58
DB 193 YRDAKGHYDEDEEFARARNYV-----VXLQSGDEYF----- 224

QY 59 SHMQORDTIANKLKDYNNHR-----DRLDVMAADDICTAINTNGEYKGLYLGPGTGTG 114
DB 225 -REMR-----KLVDTITQONQITYDRLVNTLTRD-----VMBSLYNP----- 263

QY 115 SFTLGAIANOLSKSKVR-----STIYLPF-----IRTLKGF----- 148
DB 264 --MLPGIVADIAKAGLAIVSEGAIVFDFEFKNKEGEPGVIIQKXDGGLYTTTDIACA 321

QY 149 -----KQGS----- 152
DB 322 KYRYETLHADRVLYYIDSRQHILMQAWAIVKAGYVPESVPLEHMFQMLGKDGKPFK 381

QY 153 -----PEKLRHVREANIMLDDIGAEBVTWVRDEVIQ----- 186
DB 382 TRAGTVKLADLDELERARRLVAEKNDMPADELEKLANAVGICAVKADISKNRITD 441

QY 187 -----PLHYRMVHELPTFFSSNPFSYSELEHHLIANTRDGEETKARIT 230
DB 442 YIFDWMNMLAFEGNTAPWQYATRVLSVFRKAIDEQDLAAAPVIRERERQALARLL 501

QY 231 ERYKLS-----TPYFL 242
DB 502 QFEETLTVAARBEGTPHVM 519

RESULT 10
US-09-976-165-13
Sequence 13, Application US/09976165
Patent No. US20020107383A1
GENERAL INFORMATION:
APPLICANT: FUJIMURA, TSUTOMU
APPLICANT: MATANABE, TAKESHI
APPLICANT: HORIE, MASATO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
FILE REFERENCE: Q-53599
CURRENT APPLICATION NUMBER: US/09/976,165
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/565,538
PRIOR FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR APPLICATION NUMBER: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR APPLICATION NUMBER: JP 63410/1996
PRIOR FILING DATE: 1996-03-19
PRIOR APPLICATION NUMBER: JP 69163/1997
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 389
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-165-13

Query Match
Best Local Similarity 35.1%; Score 90; DB 10; Length 389;
Pred. No. 0.35;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 102 KGLIYGPBGTSFLLGAIANOLKS---KKVRSITII--YLPEFIRTLKGFQDSFEK 156
DB 168 KGLIYGPBGTKTLARAVASQDCNFKVSSIVDKYIGSARILIRMF-----N 220

QY 157 LHRVREANILMDIGA 173
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 11

US-10-205-219-165
Sequence 165, Application US/10205219
Publication No. US20030138803A1
GENERAL INFORMATION:
APPLICANT: Wainer-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alistair
APPLICANT: Brooksbank, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018200
CURRENT FILING DATE: 2002-07-24
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 165
LENGTH: 389
TYPE: PRT
ORGANISM: *Spermophilus tridecemlineatus*
FEATURE:
OTHER INFORMATION: ATPase domain protein 44
US-10-205-219-165

Query Match
Best Local Similarity 35.1%; Pred. No. 0.35; Length 389;

Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 102 KGLIYGPPTGKSTILGAIANOLKS---KKVRSITI--YLPEFIRTLKGRDGSFEKK 156
DB 168 KGLIYGPPTGKSTILGAIANOLKS---KKVRSITI--YLPEFIRTLKGRDGSFEKK 156
QY 157 LHRVREANILMDIGA 173
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 12

US-10-342-276-13
Sequence 13, Application US/10342276
Publication No. US20030143688A1
GENERAL INFORMATION:
APPLICANT: FUJIMURA, TSUTOMU
APPLICANT: MATANABE, TAKESHI
APPLICANT: HORIE, MASAO
TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
FILE REFERENCE: 0-53599
CURRENT FILING DATE: US/10/342,276
CURRENT FILING DATE: 2003-01-15
PRIOR FILING DATE: US/09/976,165
PRIOR FILING DATE: 2001-10-15
PRIOR FILING DATE: 09/565,538
PRIOR FILING DATE: 2000-05-05
PRIOR FILING DATE: 09/055,699
PRIOR FILING DATE: 1998-04-07
PRIOR FILING DATE: 08/820,170
PRIOR FILING DATE: 1997-03-19
PRIOR FILING DATE: 1996-03-19
PRIOR FILING DATE: 1996-03-19
PRIOR FILING DATE: 1997-03-05
NUMBER OF SEQ ID NOS: 95
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 389

TYPE: PRT
ORGANISM: *Homo sapiens*
US-10-342-276-13

Query Match
Best Local Similarity 35.1%; Pred. No. 0.35; Length 389;
Matches 27; Conservative 13; Mismatches 25; Indels 12; Gaps 3;

QY 102 KGLIYGPPTGKSTILGAIANOLKS---KKVRSITI--YLPEFIRTLKGRDGSFEKK 156
DB 168 KGLIYGPPTGKSTILGAIANOLKS---KKVRSITI--YLPEFIRTLKGRDGSFEKK 156
QY 157 LHRVREANILMDIGA 173
DB 221 YARDHQPCTIFMDEIDA 237

RESULT 13

US-10-128-714-3421
Sequence 3421, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroskin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
FILE REFERENCE: 10182-018-999
CURRENT FILING DATE: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR FILING DATE: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR FILING DATE: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR FILING DATE: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR FILING DATE: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR FILING DATE: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3421
LENGTH: 393
TYPE: PRT
ORGANISM: *Aspergillus fumigatus*
US-10-128-714-3421

Query Match
Best Local Similarity 33.8%; Pred. No. 0.35; Length 393;
Matches 26; Conservative 17; Mismatches 22; Indels 12; Gaps 4;

QY 102 KGLIYGPPTGKSTILGAIANOLKS---KKVRSITI--YLPEFIRTLKGRDGSFEKK 156
DB 172 KGLIYGPPTGKSTILGAIANOLKS---KKVRSITI--YLPEFIRTLKGRDGSFEKK 156
QY 157 LHRVREANILMDIGA 173
DB 229 ---HEPCTIFMDEIDA 241

RESULT 14

US-09-815-242-5010
Sequence 5010, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel

```

1  APPLICANT: Trawick, John D.
2  APPLICANT: Carr, Grant J.
3  APPLICANT: Yamamoto, Robert T.
4  APPLICANT: Xu, H. Howard
5  TITLE OF INVENTION: Identification of Essential Genes in
6  TITLE OF INVENTION: Prokaryotes
7  FILE REFERENCE: ELITRA.01A
8  CURRENT APPLICATION NUMBER: US/09/815,242
9  CURRENT FILING DATE: 2001-03-21
10 PRIOR APPLICATION NUMBER: 60/191,078
11 PRIOR FILING DATE: 2000-03-21
12 PRIOR APPLICATION NUMBER: 60/206,848
13 PRIOR FILING DATE: 2000-05-23
14 PRIOR APPLICATION NUMBER: 60/207,727
15 PRIOR FILING DATE: 2000-05-26
16 PRIOR APPLICATION NUMBER: 60/242,578
17 PRIOR FILING DATE: 2000-10-23
18 PRIOR APPLICATION NUMBER: 60/253,625
19 PRIOR FILING DATE: 2000-11-27
20 PRIOR APPLICATION NUMBER: 60/257,931
21 PRIOR FILING DATE: 2000-12-22
22 PRIOR APPLICATION NUMBER: 60/269,308
23 PRIOR FILING DATE: 2001-02-16
24 NUMBER OF SEQ ID NOS: 14110
25 SOFTWARE: FastSeq for Windows Version 4.0
26 SEQ ID NO 5010
27 LENGTH: 445
28 TYPE: PR1
29 ORGANISM: Enterococcus faecalis
30 US-09-815-242-5010

```

```

Query Match      6.8% Score 90; DB 9; Length 445;
Best Local Similarity 22.1%; Pred. No. 0.42;
Matches 44; Conservative 30; Mismatches 55; Indels 70; Gaps 9;

QY      27  LYVNNRIKIRYLCQCKIKYD----- 48
      |:::|:::|:::|:::|:::|
Db      34  LKLNINQL---WLEVPASVHRDYEKNLSAKIVETGFKLTGAEWPHFVADKDALAQ 90
      |:::|:::|:::|:::|:::|

QY      49  --EEEPFEELTSHMQORDTLNQL-----KDIYNNHBDRLDVMAAADICATLNG 98
      |:::|:::|:::|:::|:::|

Db      91  ELEPPAESEEVFSEOSKKAMINPKYTFPTFVIYGNQAHAAAL---VVAED-----PG 141
      |:::|:::|:::|:::|:::|

QY      99  EQVKGVLVLYGPFQTSKSPILGAINQL---KSKKVR--STIIYLPFPIRTLKGGFKDGS 152
      |:::|:::|:::|:::|:::|

Db      142  SIYNPFLFYGGVGGLKTHLMHAIGHQMLVNPDAKVKVVSSETFTNEINSIQFTISE-Q 200
      |:::|:::|:::|:::|:::|

QY      153  FEKLHVRREANIIMLDI 171
      |:::|:::|:::|:::|

Db      201  FRKE---YRNVDLILVDDI 216
      |:::|:::|:::|:::|

RESULT 15
US-09-815-242-10908
; Sequence 10908, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zykkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIORITY FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848

```

```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-22
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10908
; LENGTH: 447
; TYPE: PR1
; ORGANISM: Enterococcus faecalis
US-09-815-242-10908

```

```

Query Match          6.8%; Score 90; DB 9; Length 447;
Beet Local Similarity 22.1%; Pred. No. 0.43;
Matches 44; Conservative 30; Mismatches 55; Indels 70; Gaps 9;

QY      27  LYDNNRIRKRYLQCPCKIKYD----- 48
      34  LKLDNNQD---WLEVPASVHRVDYWKMLSAKIVETGPELTGAENVPHFVADKDALAQ 90
Dh      49  --EERFASAILTSHMORDTLNALT-----DIYNNHRDRLDVMAAADICTATNG 98
      91  ELEPRAEVEVFSQSKAMLNPRYTDFTVYIGNQMAHAAL---VVAED-----PG 141
Dh      99  EGVKGLVLYGPEGTGSPFIIAANQD---KSKRRV--STIYLPFPIRTLKGFGKDG 152
QY      142  SIYNPDLFFYGVGIGKTHLMAIGHQMLVNPQDAKVVSSEFTNEFINSIQRTISE-Q 200
Dh
QY      153  FEKKLHRVREANTILMDI 171
Dh      201  FRKE---YRVNDLLVDDI 216

```

Search completed: December 15, 2003, 15:19:03
Job time : 25.4154 Secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 14.099 Seconds
(without alignments)
1705.235 Million cell updates/sec

Title: US-09-689-952-18

Perfect score: 1328
Sequence: 1 YKDQKHVDGKHFADCPNFV.....ERVKSLSTPYFLSGENFRNN 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1328	100.0	306	2 A89952	primosomal protein
2	537.5	40.5	311	1 I0BS44	primosome componen
3	531.5	40.0	311	2 H84042	primosome componen
4	475.5	35.8	307	2 AB1632	primosome componen
5	472.5	35.6	307	2 AH1269	primosome componen
6	381	28.7	298	2 B98066	primosome componen
7	361	27.2	298	2 D95199	primosomal protein
8	342	25.8	293	2 H86718	primosomal protein
9	173	13.0	209	2 S77882	dnaa protein homol
10	145	10.9	282	2 D97138	DNA replication pr
11	140	10.5	316	2 B82907	conserved hypothe
12	138.5	10.4	235	2 E70378	DNA replication pr
13	127.5	9.6	313	2 D69945	phage-related proc
14	127	9.6	266	2 140411	PBSX prophage ORF
15	125.5	9.5	440	2 D81415	chromosomal repli
16	120.5	9.1	316	2 T10436	probable transposa
17	115.5	8.7	246	2 D84932	DNA replication pr
18	115	8.7	399	2 G70328	chromosome replica
19	114	8.6	265	1 BVEIC7	16Sb protein - Bac
20	112	8.4	287	2 AD1588	similitarities staph
21	112	8.4	287	2 AG1651	similitarities staph
22	112	8.4	291	2 A86681	hypothetical prote
23	112	8.4	520	2 H69125	hypothetical prote
24	111.5	8.4	475	2 C96503	protein FecI6.7 i
25	111	8.4	350	2 B97735	probable ATPase n2
26	109.5	8.2	329	2 D97340	DNA replication pr
27	108.5	8.2	329	2 D97340	AAA family ATPase
28	107	8.1	585	2 P90460	hypothetical prote
29	106.5	8.0	251	2 S23889	hypothetical prote

30	106.5	8.0	630	2 AC1309	probable ABC trans
31	106	8.0	798	2 B71196	probable transactio
32	105.5	7.9	248	2 C85626	probable DNA repli
33	105.5	7.9	248	2 B90763	probable DNA repli
34	105.5	7.9	699	2 B72479	probable transactio
35	105	7.9	246	2 C90913	probable DNA repli
36	105	7.9	248	1 C64886	DNA replication pr
37	105	7.9	795	2 F75154	cell division cont
38	104.5	7.9	248	2 A85747	probable DNA repli
39	104.5	7.9	248	2 A99872	probable DNA repli
40	104.5	7.9	382	2 H72255	myo-inositol-1-pho
41	104.5	7.9	454	2 D84931	chromosomal repli
42	104	7.8	248	2 AC0738	probable DNA repli
43	104	7.8	454	1 IQVJBA	replication initia
44	103.5	7.8	263	2 T47074	hypothetical prote
45	103.5	7.8	726	2 D72613	probable transactio

ALIGNMENTS

```
RESULT 1
A89952
primosomal protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: A89952
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; I
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89952
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-306 <KUR>
A:Cross-references: GB:BA000018; PID:g13701480; PIDN:BAE42774.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: dnaI
A:Superfamily: 44K dnaa protein homolog

Query Match      100.0%; Score 1328; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 1.3e-100;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKDQKHVDGKHFADCPNFVKGVPELVYDNNRIKIRYLQCPCKIKYDERFEALITSH 60
    |||
DB 57 YKDQKHVDGKHFADCPNFVKGVPELVYDNNRIKIRYLQCPCKIKYDERFEALITSH 116
    |||

QY 61 HMQDPTLNAKLDIYNNHNRDLVAMADDTCTAITNGEQVGLYLYGPGFGKSFILGA 120
    |||
DB 117 HMQDPTLNAKLDIYNNHNRDLVAMADDTCTAITNGEQVGLYLYGPGFGKSFILGA 176
    |||

QY 121 INQKSKKVRSTIYLPFRITLKGFGDGSFEKKHVRANITIMDDIABEYTPV 180
    |||
DB 177 INQKSKKVRSTIYLPFRITLKGFGDGSFEKKHVRANITIMDDIABEYTPV 236
    |||

QY 181 RDEVIGPLHYMVHMLPTFFSSNFDYSELHHLAMTRDGESEKTYAARIIEVKSLSTPY 240
    |||
DB 237 RDEVIGPLHYMVHMLPTFFSSNFDYSELHHLAMTRDGESEKTYAARIIEVKSLSTPY 296
    |||

QY 241 FLSGENFRNN 250
    |||
DB 297 FLSGENFRNN 306
    |||

RESULT 2
I0BS44
primosome component (helicase loader) dnaI - Bacillus subtilis
N:Alternate names: dnaa protein homolog, 44K; hypothetical protein Y (dnaB 3' region)
C:Species: Bacillus subtilis
```

C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 19-Jan-2001
C/Accession: B24720; C26580; F69617
R/Organism: N. Motiia, S. Mazza, P.G.; Yoshikawa, H.
Nucleic Acids Res. 14, 9989-9999, 1986
A/Title: Nucleotide sequence and organization of dnaB gene and neighbouring genes on the
A/Reference number: A93650; MUID:87117549; PMID:3027671
A/Accession: B24720
A/Molecule type: DNA
A/Residues: 1331 <GGA>
A/Cross-references: GB:X04963; NID:939880; PIDN:CAA28633.1; PID:G39881
R/Hoshino, T.; McKenzie, T.; Schmidt, S.; Tanaka, T.; Sueoka, N.
Proc. Natl. Acad. Sci. U.S.A. 84, 653-657, 1987
A/Title: Nucleotide sequence of Bacillus subtilis dnaB: a gene essential for DNA replica
A/Reference number: A94709; MUID:87118226; PMID:3027697
A/Accession: C26580
A/Molecule type: DNA
A/Residues: 1-18, 'N', 20-23, 'T', 25-206 <HOS>
R/Kunze, P.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fujita, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gall
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holappel, S.; Hosono, S.; Hullo, M.
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A/Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanton
A/Authors: Schleicher, S.; Schroeter, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumsen, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: F69617
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-311 <KUN>
A/Cross-references: GB:Z99118; GB:AL009126; NID:92635200; PIDN:CAB14858.1; PID:92635363
A/Experimental source: strain 168
C/Genetics:
A/Gene: dnaI
C/Superfamily: 44k dnaa protein homolog
C/Keywords: ATP; nucleotide binding; P-loop
F/168-175/Region: nucleotide-binding motif A (P-loop)
Query Match 40.5%; Score 537.5; DB 1; Length 311;
Best Local Similarity 44.9%; Pred. No. 2.8e-36;
Matches 106; Conservative 48; Mismatches 81; Indels 1; Gaps 1;
QY 15 DCPNPKVGHPELVYNNRIRIKIRYLQCPCKIKYDEBFPEALITSHHMDTLNAKLDI 74
DB 75 NCNNLEIGHPKVYNNRISIDIEYECVPRKADQCKQKSLKSMITQDLDGATFQV 134
QY 75 YNNHRDLVAMAAADICTAITNGEOVKGLYVPGFGTSGFTLIGAIANOLSKKVRSTI 134
DB 135 DISDPRLAMFGVHTVPLKSYNETGKGLYLVGKGVGTFWLAIANLAKKXSSWI 194
QY 135 TYPEPRTLKGGFGKGSFFKKLHRYREANITMLDGAEEVTPWRDEVYGLLHYRWY 194
DB 195 VVYVEFVRELKNSLQDQTLLEKLNMYKTPVLMLDIGASMTSWRDEVYGLVLRMS 254
QY 195 HELPTFFSSNFDSSELEHNLAMTRDGE-EKTKAARIERKSLSTPFLSGENFRN 249
DB 255 QQLPTFFSSNFSPDELDGHTFTYSGRGEKEVKARLMERILYLAAPFLRDEGNRRH 310

Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliophilic bacterium Bacillus halodurans and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: H84042
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-311 <STO>
A/Cross-references: GB:AP001517; GB:BA000004; NID:910175500; PIDN:BA006863.1; GSPDB:GN001
A/Experimental source: strain C-125
C/Genetics:
A/Gene: dnaI
C/Superfamily: 44k dnaa protein homolog
Query Match 40.0%; Score 531.5; DB 2; Length 311;
Best Local Similarity 41.2%; Pred. No. 8.5e-36;
Matches 107; Conservative 55; Mismatches 79; Indels 19; Gaps 6;
QY 1 YKQDQKHVDGH--KPADCPNFKVGHPELVYNNRIRIKIRYLQCPCKIKYDEBFPEALIT 58
DB 59 YKQDQSHC-AHCPGLQCKNLMKGYPTLYVERDSLSYSPCKKEERREKKRSLR 117
QY 59 SHHMDTLNAKLDI-----YNNHRDLVAMAAADICTAITNGEOVKGLYVPGFG 111
DB 118 SYTIPKEILEAFDVSERPSISHR-ALBFALSAKP-----GEGMGILYKDG 169
QY 112 TGKSPFLGAIANOLSKKVRSTIYPERIRTLKGGFGKGLHRYREANITMLD 171
DB 170 VGTFTPLMGAIANLCKRGIDSTIVVDPDFRELKOSIGDTQCKDLYKNAQVLLFD 229
QY 172 GAEVTPWRDEVYGLLHYRWVHELPTFFSSNFDSLEHNLAMT-RDGEKTKAARI 230
DB 230 GAEVTPWRDEVYGLLHYRWVHELPTFFSSNFDSLEHNLAMT-RDGEKTKAARI 289
QY 231 ERYKSLSTPFLSGENFRN 250
DB 290 ERIHYTVSVMWQGNRRH 309

RESULT 4
AB1632
primosome component (helicase loader) dnaI [imported] - Listeria innocua (strain C1p112)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: AB1632
R/Glaeser, P.; Frangoul, L.; Buchliesser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesunger, O.; Entian, K.D.; Feigl, H.;
Science 294, 849-852, 2001
A/Authors: Krefc, J.; Kuhn, M.; Kunst, F.; Kurapkac, G.; Madueno, E.; Maitournam, A.; Mat
ok, C.; Schlueder, T.; Simoes, N.; Tillerer, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1632
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-307 <GLA>
A/Cross-references: GB:AL92022; PIDN:CMC96826.1; PID:916414082; GSPDB:GN00178
A/Experimental source: strain C1p11262
C/Genetics:
A/Gene: dnaI
C/Superfamily: 44k dnaa protein homolog
Query Match 35.8%; Score 475.5; DB 2; Length 307;
Best Local Similarity 42.6%; Pred. No. 3e-31;
Matches 98; Conservative 45; Mismatches 84; Indels 3; Gaps 2;
QY 22 GHVPELVYNNRIRIKIRYLQCPCKIKYDEBFPEALITSHHMDTLNAKLDIYNNHRDR 81
DB 77 GYAPKVLNGBERTIYTYTKKIEEDKRAVERRRISLYMKQVVDNALADPYTDEBR 136
QY 82 LDVAMAAADICTAIT--TNGEOVKGLYVPGFGTSGFTLIGAIANOLSKKVRSTIY 139
DB 137 KVALVEAYGFLANNYPPOGNERVKGFLTHOSFGKGSYLLGALAKELALNGISTVL 196

[illegible]

C:Genetics: dnal
 A:Gene: dnal
 C:Superfamily: 44k dnaa protein homolog
 Query Match 28.7%; Score 381; DB 2; Length 298;
 Best Local Similarity 34.7%; Pred. No. 1,4e-23;
 Matches 82; Conservative 50; Mismatches 100; Indels 4; Gaps 3;
 QY 15 DCPNFKVGHVPELVYDNNRIKIRYLQCPCKIKYDERFPAELITSHHMRDITLNAKLDI 74
 DB 65 DTDYIAKQKPELVVNNHGVADVSYEETPELIAEKAIAIKRNLINPSSLKAVSFVLDV 124
 QY 75 YNNHNRDLRDVMAADICTAINTGEQVKGILYGPFGTGKSFILGAIANOLSKKK-V 130
 DB 125 YRDVQRLLTLKRMIEFVNDYFN--NLKGLYIGDVGKGSFVVAALAHDLSEKRGVSST 182
 QY 134 IIVPEFIRTLKGFCKDGSFEKKLHVRREANIIMLDDIGAEEVTPWVRDEVYIGPLHYRM 193
 DB 183 LLHYSPFIDVGNALSDGNVKTLYDEIKLSEVILDDIGAEGSTVWVRDELQVILQYRM 242
 QY 194 VHELPFPSSNFDYSLEHHLAMTRDGEKTKAANIIEKVLSTPYPLSGENFR 248
 DB 243 QENLPFTFSNPFEDLEKGFPAKVKHGNDETWEARRVERIRYLAETRETLGAVNR 298
 RESULT 7
 D95199
 Primosomal protein Dnal [imported] - Streptococcus pneumoniae (strain TIGR4)
 C:Species: Streptococcus pneumoniae
 C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C:Accession: D95199
 R:Retcelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
 son, T.; Hickey, B.K.; Holt, I.B.
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; M01D:21357209; PMID:11463916
 A:Accession: D95199
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-298 <K1B>
 A:Cross-references: GB:AE005672; P1DN:AAK75789.1; PID:g14973206; GSPDB:GN00164; TIGR:SP4
 A:Experimental source: strain TIGR4
 C:Genetics:
 A:Gene: SP1711
 C:Superfamily: 44k dnaa protein homolog
 Query Match 27.2%; Score 361; DB 2; Length 298;
 Best Local Similarity 34.7%; Pred. No. 6e-22;
 Matches 83; Conservative 50; Mismatches 96; Indels 10; Gaps 4;
 QY 15 DCPNFKVGHVPELVYDNNRIKIRYLQCPCKIKYDERFPAELITSHHMRDITLNAKLDI 74
 DB 65 DTDYIAKQKPELVVNNHGVADVSYEETPELIAEKAIAIKRNLINPSSLKAVSLAQV 124
 QY 75 YNNHNRDLRDVMAADICTAINTGEQVKGILYGPFGTGKSFILGAIANOLSKKK-V 130
 DB 125 DLDGGRPLVPEFKLAFVQYPAIR----KGLYIGDVGKGSFVVAALAHDLSEKRGV 179
 QY 131 RSTIIVPEFIRTLKGFCKDGSFEKKLHVRREANIIMLDDIGAEEVTPWVRDEVYIGPLH 190
 DB 180 SSTLLHYSPFIDVGNALSDGNVKTLYDEIKLSEVILDDIGAEGSTVWVRDELQVILQ 239
 QY 191 YRMVHELPFPSSNFDYSLEHHLAMTRDGEKTKAANIIEKVLSTPYPLSGENFR 248
 DB 240 YRMQENLPFTFSNPFEDLEKGFPAKVKHGNDETWEARRVERIRYLAETRETLGAVNR 298
 RESULT 8
 H86718
 Primosomal protein Dnal [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis

C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C/Accession: H86718
 R/Polioin. A.; Wincer, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Enria
 Genome Res. 11, 731-753, 2001
 A/Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s8
 A/Reference number: A86625; M01D:21235186; PMID:11337471
 A/Accession: H86718
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1293 <SNO>
 A/Cross-references: GB:AE005176; PID:g12723668; PIN:AAK04850.1; GSPDB:GN00146
 A/Experimental source: strain IL1403
 C/Genetic:
 A/Genetic: dna1
 C/Superfamily: 44k dnaa protein homolog

Query Match 25.8%; Score 342; DB 2; Length 293;
 Best Local Similarity 33.3%; Pred. No. 2.1e-20;
 Matches 85; Conservative 49; Mismatches 97; Indels 24; Gaps 6;

QY 2 KDOQKHVDGKHPDCRFVAGVDELVDNNRIRKIRLQCPCKIKYDEERFEALITSH 61
 DB 57 REHKEKKEKRA---ADGEVLIMNHGVADVST---QTNELAQQAQNLRRMN 108
 QY 62 ---MORDTLNAKLDIYNHNRDLVDVMAADICTAITNGEQVGLYLPFGTGSFIL 118
 DB 109 IILGPDOLKQVTLADILDVQRIKPYQALVDF---ITNPPKKGILYLDVFGVGSFML 165
 QY 119 GAIANOLKSKKVRSTIYLDPEFIRTLKGFQDGSFEKK--LHRVRANITMLDGAEE 175
 DB 166 AAMANEILAKKIGISTITLLHYPTFI-----SDLDFNAKRWVNEIKASQVLVLDIGAEG 218
 QY 176 VTFWVDEVIGPLHYRMVHELPTFESSNEDYSELBNHLMTRDGEKTKAARLIERVKS 235
 DB 219 NNAWVRDSTIQVTLQHRMGOENLPTFTSNLRMBELQHLAETKRADEIWPAAKVMEXKY 278
 QY 236 LSTPYPLSGENFRNN 250
 DB 279 LAEMRLLEGTRRRHD 293

RESULT 9

77882
 dnaa protein homolog - Mycoplasma capricolum
 N/Alternate names: protein MC072
 C/Species: Mycoplasma capricolum
 C/Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 07-Dec-1999
 C/Accession: S77882
 R/Bork, P.; Ouzounis, C.; Casati, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.;
 Mol. Microbiol. 16, 955-967, 1995
 A/Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiolo
 A/Reference number: S77739; M01D:96059641; PMID:7476192
 A/Accession: S77882
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-209 <BOR>
 A/Cross-references: EMBL:Z33058; NID:g514450; PIDN:CA83732.1; PID:g530419
 A/Experimental source: ATCC 27343
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
 C/Genetic:
 A/Genetic code: SGC3

Query Match 13.0%; Score 173; DB 2; Length 209;
 Best Local Similarity 31.9%; Pred. No. 7.4e-07;
 Matches 51; Conservative 31; Mismatches 58; Indels 20; Gaps 8;

QY 45 IKYDERFEALITSHMORDTLNAKLDIYNHNRDLVDVMAADICTAITNGEQVGL 104
 DB 57 LTLD-EYFKNEL---NNSESLDQKEDI---RKL-----LSNLRKVQNSK-KGV 102
 QY 105 YLYGPGTGKSFILGAIANOLKSKKVRSTIY--LPEFIRTLKGFQDGSFE--KKLHRV 160
 DB 103 YLYGSGIGKTYMFKVLANTLASK--NKTVIFSTLRSLIDKLKESFNSSEINSELEMKKI 160

QY 161 REANITMLDIGAEEVTPWVRDEVIGELLYHVMHLPTE 200
 DB 161 KTYVDFLPLDDIGGENISLWARDLFEVLNRYMKNKATF 200

RESULT 10

D97138
 DNA replication protein DnaC [imported] - Clostridium acetobutylicum
 C/Species: Clostridium acetobutylicum
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C/Accession: D97138
 R/Molling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
 J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
 A/Reference number: A96900; M01D:21359325; PMID:21359325
 A/Accession: D97138
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-282 <KUR>
 A/Cross-references: GB:AE001437; PIDN:AAK79895.1; PID:g15024913; GSPDB:GN00168
 A/Experimental source: Clostridium acetobutylicum ATCC824
 C/Genetic:
 A/Genetic: CAC1933

Query Match 10.9%; Score 145; DB 2; Length 282;
 Best Local Similarity 21.8%; Pred. No. 0.00021;
 Matches 51; Conservative 52; Mismatches 103; Indels 28; Gaps 7;

QY 32 NRIRIKYVLOPCPKIKYDERFEA-----ELITSHMORDTLNAKLDIYNHNRDL 82
 DB 44 DRIVKTMCEESAVMHEKQHEMLKKQNRNELITNSLMDKFRSNEENMFNGSR 103
 QY 83 DVAMA---ADDI--CTAITNGEQVGLYLYGPGTGKSFILGAIANOLKSKKVRSTIY 137
 DB 104 MYNIGRRVADNFECK-----QQGLDILYGSXGKNTYLAISANIELKQYIPVVCVS 158
 QY 138 PERIRTLKGFQDGSFEKKLHRV---EANIITMLDIGAEEVTPWVRDEVIGELLYHVM 194
 DB 159 NGLISRIQKTYNSWGEASDITIRSFINDLITDGLTEKSEWSR-SNITNIVSR 217
 QY 195 HELPTFSSNPDYSLEHNLAMTRDGEKTKAARLIERVKSISTPYPLSGENFR 248
 DB 218 SKPLITITNLEINPSKRHGVADQYHERTES-----RIFEMOTPVENTSISR 266

RESULT 11

B82907
 conserved hypothetical ATP/GTP-binding protein U0307 [imported] - Ureaplasma urealyticum
 C/Species: Ureaplasma urealyticum
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C/Accession: B82907
 R/Glaes, J.I.; Lefkowitz, E.J.; Glaes, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
 submitted to Genbank, February 2000
 A/Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mini
 A/Reference number: A82870
 A/Accession: B82907
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-316 <GLA>
 A/Cross-references: GB:AE002128; GB:AF222894; NID:g6699279; PIDN:AAF30716.1; GSPDB:GN0015
 A/Experimental source: serovar 3, biovar 1
 C/Genetic:
 A/Genetic: U0307
 A/Genetic code: SGC3

Query Match 10.5%; Score 140; DB 2; Length 316;
 Best Local Similarity 25.5%; Pred. No. 0.00062;
 Matches 39; Conservative 28; Mismatches 68; Indels 18; Gaps 3;

QY 97 NEQVQGLVLYGPGTGKSFILGAIANOLKSKKVRSTIYLDPEFIRTLKGFK----- 149
 DB 103 YLYGSGIGKTYMFKVLANTLASK--NKTVIFSTLRSLIDKLKESFNSSEINSELEMKKI 160

A:Gene: ydam

C:Genetics:

A:Experimental source: strain 168

A:Cross-references: GB:299117; GB:AL009126; NID:g2634966; PIDN:CAB14567.1; PID:e1183855

A:Molecule type: DNA

A:Residues: 1-313 <KUN>

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Accession: D69945

A:Reference number: A69580; MUID:98044033; PMID:9384377

T.: Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, T.; Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A69580; MUID:98044033; PMID:9384377

D: 217 AKLDAMRKVEVLFIDDLFKPINOQPRATDW-QVEQIOSVLNRYLRANKPLLISSSELTIDE 275

OY 210 L 210

DB 276 I 276

Query Match
Best Local Similarity 33.1%; Pred. No. 0.0063;
Matches 40; Conservative 19; Mismatches 49; Indels 13; Gaps 6;

OY 98 GEQVGKLTVGPPTGKSFLIGAIANOL-KSKYRSTIITLP--EFRRTLKGFQDSFE 154
|||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 161 GERONSTALIGQGSGSKTHLLTAIMNNLIIKKSVH--CMTFPYVEGMGDIAKF--DNLE 216
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
OY 155 KLHRVRANIILMDLI-----GAERTVPWRREVIGPLLHYRMVHELPFFPSNPFSYE 209
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
DB 217 AKLDAMRKVEVLFIDDLFKPINOQPRATDW-QVEQIOSVLNRYLRANKPLLISSSELTIDE 275

RESULT 14

140411
PBsx prophage ORF xkdc - Bacillus subtilis

N;Alternate names: protein 5 (xre region)

C;SpecDate: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999

A;Accession: 140411 | C69731 | S47111

J;McDonnell, G.E., Wood, H., Devine, K.M.; McConnell, D.J.

J;Bacteriol. 176, 5820-5830, 1994

A>Title: Genetic control of bacteriophage induction: regulation of the induction of PBsx in B.

A;Reference number: 140408; PMID:94364963; PMID:8083174

A;Accession: 140411

A>Status: preliminary; translated from GB/EMBL/DBD/J

A:Molecule type: DNA

A;Residues: 1-266 <RES>

A;Cross-references: EMBL:Z34287; NID:G498810; PIDN:CA84044.1; PID:G498814

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Avevedo, V.; Bertero, V.; Brocchi, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chakrabarti, S.D.; Emmertson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Gallerich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidos, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Matsuda, R.; Ogawa, K.; Ogiwara, A.; Ondega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Poterrelli, R.; Rieger, M.; Rivolta, C.; Kocha, E.; Roche, B.; Rose, M.; Saaipe, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scrofione, F.; Sekiguchi, J.; Sekowska, A.; Serotoni, A.; Takanaka, M.; Tanekoshi, A.; Tanaka, T.; Terpilstra, P.; Toiguchi, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, H.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: C69731

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A;Residues: 1-238, 'K', 240-266 <XDN>

A;Cross-references: GB:F599110; GB:A10093126; NID:g2633472; PIDN:CAB13110.1; PID:e1183273

A;Experimental source: strain 168

C;Genetics:

A;Gene: xkdc

Query Match 9.6%; Score 127 DB 2; Length 266;

Best Local Similarity 31.0%; Pred. No. 0.0056;
Matches 39; Conservative 25; Mismatches 44; Indels 18; Gaps 7;

QY CTA--ITNGEYQR-----GLVYGFPGTAKSFLIGAIQAQSKSKVSTIYLP--EFLR 142
Db CTKEVAADYEQIKDCKRKSIALIGQPGSGKTHLFFAAHEL-MRTCYVPVIYPPFVEGFT 159
QY TLKGGFGKGSFPKCLHVRSEANIIMDDI-----GAEVTPVWRDVIQGLIHYRWHEL 197
Db DLKNDP--ALLBAKLNRMKQADVLFIIDFLFKPVNGKPRATDQMLEGMY-S-VLNRYRLNKH 216
QY PTFFSS 203
Db PILSS 222

RESULT 15

chromosomal replication initiation protein Cj0001 [imported] - Campylobacter jejuni (strain
 C/Spectes: Campylobacter jejuni
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
 C/Accession: D81415
 R./Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Chilling
 C.W.; Quail, M.; Rajandream, M.A.; Kutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell
 Nature 403, 665-668, 2000
 A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
 A/Reference number: A81250; MUID:20150912; PMID:10688204
 A/Accession: D81415
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-440. <PAR>
 A/Cross-references: GB:AL139074; GB:AL111168; NID:66967505; PIDN:CBM72494.1; PID:g656750
 A/Experimental source: serotype O2, strain NCTC 11168
 C/Genetics:
 A/Gene: dnaA, Cj0001
 A/Superfamily: replication initiation protein dnaA

Query Match	9.5%;	Score 125.5;	DB 2;	Length 440;
Best Local Similarity	26.8%;	Pred. NO. 0.014;		
Matches 41;	Conservative 30;	Mismatches 51;	Indels 31;	Gaps 7;

Qy 92 CTAITNGEONKGY-----LYGPGGKSPFIIAGINQKSKVSTIYL--PEPIITLK 145
 Db 123 CKRIAHKDKGLKLYNPIFYVGPPTGKTHLLQAVGN--ASLEMGKVITYATSENFINDP 180
 Qy 146 GGPFGSGFKLHRYEANIILMDIDIGAEVYLPWVRDEVIGPLLHMYMHELPTFFESNF 205
 Db 181 SNKNGSLDKFHEKTRNCVDVLIIDVQVGLGTDKIQEE-----FF---F 221
 Qy 206 DYSELEHH--LAAFRDGEKTKAARIYERYS 235
 Db 222 IFNEIKNNDQIILMTSDNPNNMLKG-ITERLKS 253

Search completed: December 15, 2003, 15:16:50
Job time : 15.099 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 7.90922 Seconds

(without alignments)
1486.453 Million cell updates/sec

Title: US-09-689-952-18
Sequence: 1 YK0QKHVDGKHFADCPNFV.....ERVKSLSPTPYFLSGENFRNN 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	537.5	40.5	311	1	DNAI_BACSU
2	128.5	9.7	457	1	DNAI_BUCBP
3	127.5	9.6	313	1	YOAM_BACSU
4	127	9.6	266	1	XKOC_BACSU
5	125.5	9.5	440	1	DNAI_CAMUJ
6	122	9.2	246	1	DNAI_BUCAP
7	115.5	8.7	246	1	DNAI_BUCAP
8	115	8.7	399	1	DNAI_AQUAE
9	114	8.6	265	1	ISTB_PSEAE
10	111.5	8.4	445	1	DNAI_ACHLA
11	109.5	8.2	229	1	ISTB_BURCE
12	107	8.1	451	1	DNAI_PASMU
13	106.5	8.0	251	1	ISTB_PASMU
14	105.5	7.9	231	1	ISTB_PSEFL
15	105	7.9	248	1	YDAP_ECOLI
16	104.5	7.9	454	1	DNAI_BUCAI
17	104	7.8	454	1	DNAI_BUCAP
18	103.5	7.8	263	1	Y4BM_RHISN
19	102.5	7.7	432	1	SKDI_SCHPO
20	102.5	7.7	466	1	DNAI_PSEMI
21	102	7.7	462	1	DNAI_PSEMI
22	102	7.7	464	1	DNAI_SERMA
23	100	7.5	263	1	ISTB_BACFR
24	100	7.5	307	1	RUVA_MYCGE
25	100	7.5	643	1	DNAI_STRRE
26	100	7.5	656	1	DNAI_STRCO
27	99.5	7.5	794	1	SC18_CANAL
28	99	7.5	4594	1	DYHC_HUMAN
29	99	7.5	4644	1	DYHC_MOUSE
30	99	7.5	4644	1	DYHC_RAT
31	98.5	7.4	453	1	DNAI_STRPN
32	98	7.4	245	1	DNAI_ECOLI
33	98	7.4	466	1	DNAI_SALTI

34	98	7.4	466	1	DNAI_SALTY
35	98	7.4	467	1	DNAI_ECO57
36	98	7.4	467	1	DNAI_ECOLI
37	98	7.4	468	1	DNAI_VIBPA
38	97.5	7.3	449	1	DNAI_BACHD
39	97.5	7.3	468	1	DNAI_VIBVU
40	97	7.3	340	1	RFC5_HUMAN
41	97	7.3	518	1	DNAI_NEIMB
42	97	7.3	2153	1	POIG_HRV16
43	96.5	7.3	440	1	PR54_CHICK
44	96	7.2	410	1	PSMR_METTH
45	96	7.2	436	1	PSMR_METTXA

ALIGNMENTS

RESULT 1	ID	DNAI_BACSU	STANDARD;	PRT;	311 AA.
AC	P06567;				
DT	01-JAN-1988 (Rel. 06, Created)				
DT	01-JAN-1988 (Rel. 06, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Primosomal protein dnaI.				
GN	DNAI.				
OS	Bacillus subtilis.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_Taxid=1423;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87117549; PubMed=3027671;				
RA	Ogasawara N., Moriya S., Maza P.G., Yoshikawa H.;				
RT	"Nucleotide sequence and organization of dnaB gene and neighbouring				
RT	genes on the Bacillus subtilis chromosome.";				
RL	Nucleic Acids Res. 14:9989-9999(1986).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	STRAIN=168;				
RX	MEDLINE=97124191; PubMed=8969504;				
RA	Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,				
RA	Sanders J., Emerson P.T., Harwood C.R.;				
RT	"The dnaB-pheA (256 degrees-240 degrees) region of the Bacillus				
RT	subtilis chromosome containing genes responsible for stress				
RT	responses, the utilization of plant cell walls and primary				
RT	metabolism.";				
RL	Microbiology 142:3067-3078(1996).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=168;				
RX	MEDLINE=98048467; PubMed=9387221;				
RA	Lapidis A., Galleron N., Sorokin A., Ehrlich S.D.;				
RT	"Sequencing and functional annotation of the Bacillus subtilis genes				
RT	in the 200 kb rmb-dnaB region.";				
RL	Microbiology 143:3431-3441(1997).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=168;				
RX	MEDLINE=98044033; PubMed=9384377;				
RA	Kuntz F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,				
RA	Azeredo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,				
RA	Borriero R., Bortner L., Brans A., Braun M., Brignell S.C., Bron S.,				
RA	Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,				
RA	Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,				
RA	Deutscher F., Devigne K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,				
RA	Dentant K.D., Errington J., Fabret C., Ferrari B., Foulger D.,				
RA	Fritz C., Fujita M., Fujita Y., Fuma S., Gallizon A., Galleron N.,				
RA	Ghim S.Y., Glaser P., Goffeau A., Golightly E.Z., Grandi G.,				
RA	Ginsberg G., Guy B.J., Haga K., Haled J., Harwood C.R., Henaut A.,				
RA	Hilbert H., Holsappel S., Hosono S., Hullo M.F., Ilaya M., Jones L.,				
RA	Johns B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Klein C.,				
RA	Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,				
RA	Kurita K., Lapidis A., Lardinois S., Lauber J., Lazarevic V.,				

RA Lee S.M., Levine A., Liu H., Maeda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Neome D., O'Reilly M., Ogawa M., Ogiwara S., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portolillo S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche M., Rose M., Sadleir Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorochin A., Tacconi B., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpetri P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandendol M., Vamler P., Vassartel A.,
RA Vairi A., Wambut R., Wedler B., Wedler H., Weitzengruber T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis* ";
RL Nature 390:249-256(1997).
RN [5]
RP SEQUENCE OF 1-206 FROM N.A.
RX MEDLINE=67118226; PubMed=3027697;
RA Hoshino T., McKenzie T., Schmidt S., Tanaka T., Sueoka N.,
RT "Nucleotide sequence of *Bacillus subtilis* dnaB: a gene essential for
RT DNA replication initiation and membrane attachment";
RL Proc. Natl. Acad. Sci. U.S.A. 84:653-657(1987).
RN [6]
RP SIMILARITY TO DNA.
RX MEDLINE=92195821; PubMed=1549481;
RA Koonin E.V., "Archaeobacterial virus SSV1 encodes a putative DnaA-like protein";
RT Nucleic Acids Res. 20:1143-1143(1992).
RN [7]
RP IDENTIFICATION.
RX MEDLINE=95291463; PubMed=7773414;
RA Bruand C., Ehrlich S.D.,
RT "The *Bacillus subtilis* dnaI gene is part of the dnaB operon";
RL Microbiology 141:1199-1200(1995).
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: SOME, TO DNAA FROM VARIOUS BACTERIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04963; CA28633.1; -;
DR EMBL; J75208; CA99605.1; -;
DR EMBL; AF008220; AAC00359.1; -;
DR EMBL; Z99118; CAB14858.1; -;
DR EMBL; M15183; AAA22405.1; -;
DR FTR; B24720; I08544.
DR Subtilin; BG10359; dnaI.
DR InterPro; IPR003593; AAA_ATPase.
DR SMART; SMO0382; AAA; 1.
DR KMW Primosome; DNA replication; ATP-binding; Complete proteome.
FT NP_BIND 168 175 ATP (PROBABLE).
FT CONFLICT 19 19 K -> N (IN REF. 5).
FT CONFLICT 24 24 M -> T (IN REF. 5).
SQ SEQUENCE 311 AA; 36114 MW; A86FC94AB6841264 CRC64;
Query Match 40.5%; Score 537.5; DB 1; Length 311;
Best Local Similarity 44.9%; Pred. No. 3.3e-37;
Matches 106; Conservative 46; Mismatches 81; Indels 1; Gaps 1;

QY 135 IYVPEFIRTLKGGFKDGSPEKTLHVRVANIIMLDDIGAEVTPWVRDEVIGPLHYRMV 194
DB 195 VYVEFVRRLKSLNLOQTLEKLNMYKTPVLMDDIGESMTSWRDEVIGVLRHMS 254
QY 195 HELPTFSSNFSDSELEHLMATRDGE-EKTKAARIERYKSISTPYPLSGENFRN 249
DB 255 QQLPTFFSSNFSDSELEHLMATRDGE-EKTKAARIERYKSISTPYPLSGENFRN 310
RESULT 2
ID DNAA_BUCBP STANDARD; PRT; 457 AA.
AC P59567;
DT 15-SRP-2003 (Rel. 42, Created)
DT 15-SRP-2003 (Rel. 42, Last sequence update)
DE 15-SRP-2003 (Rel. 42, Last annotation update)
DE Chromosomal replication initiator protein dnaA.
GN DNAA OR BBP012.
OS Buchnera aphidicola (subsp. *Batzongia pisicifera*).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva P.J.,
RA Tamames J., Vignera E., Latorre A., Valencia A., Moran F., Moya A.,
RT "Reductive genome evolution in *Buchnera aphidicola*.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -1- FUNCTION: Plays an important role in the initiation and regulation
CC of chromosomal replication. Binds to the origin of replication; it
CC binds specifically to double-stranded DNA at a 9 bp consensus (dnaA
CC box): 5'-TTATC(C/A)A(C/A)-3'. DnaA binds to ATP and to acidic
CC phospholipids (By similarity).
CC -1- SIMILARITY: Belongs to the dnaA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE014016; AA026756.1; -;
DR HAMAP; MF 00377; -; 1.
DR SMART; SMO0382; AAA; 1.
DR PROSITE; PS01008; DNAA; FALSE NEG.
DR KMW DNA replication; DNA-binding; ATP (POTENTIAL).
FT NP_BIND 162 169 ATP (POTENTIAL).
SQ SEQUENCE 457 AA; 53365 MW; B2950813FADD4028 CRC64;
Query Match 9.7%; Score 128.5; DB 1; Length 457;
Best Local Similarity 22.3%; Pred. No. 0.002;
Matches 67; Conservative 46; Mismatches 105; Indels 83; Gaps 12;
QY 21 KGHVPELYVDN---NRIRYQ-----CP-----CKIYDEERPEAL 56
DB 34 KNHILLYVAPNERSFWIKDYENLKLKNEFCNINTPTMLKICKPKIIOKFFNEL 93
QY 57 IYSHHQRDTLANKADIMNHRDLVMAADDICTATNGEQVY----- 103
DB 94 T---LKNILNSKL--YVNTKLSNIYSSSEINTVTFONFTGQSNOLAFTIYKI 146
QY 104 -----LYLGPFTGSKFIIAIALNOLSKKRVSTIYL--PEFIRTLKGGFKD 150
DB 147 AHNPGKYNPNPLFLYKSGSLGKTHLHVAANTLKTKNTKIYINSENFQIMTSLKN 206
QY 151 GSFPEKTLHVRVANIIMLDDIGAEVTPWVRDEV---IGPLHY-----RMVH 195
DB 207 NTIEEFKYYRSVNTLLIDIOFPAVKKHQSQELFTFTIALLARNQOIIITSDQFOKIH 266

QY 196 ELPTFSSNDYSELHHLMTDRGEKTKAARI---ERKSLSTPY---PLSGENFRN 249
 DB 267 GIFTRLKSRF-----ECGLTIRIDPPDLNTRTKILKKSHIYDINLSYKAAFLPAKLKLS 321
 QY 250 N 250
 DB 322 N 322

RESULT 3
 YQM_BACSU
 ID YQM_BACSU STANDARD; PRT; 313 AA.
 AC P45910;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein yqam.
 GN YQAM.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=95219086; PubMed=7704261;
 RA Takemaru K.-I., Mizuno M., Sato T., Takeuchi M., Kobayashi Y.,
 RT "Complete nucleotide sequence of a skin element excised by DNA
 rearrangement during sporulation in *Bacillus subtilis*.",
 RL Microbiology 141:323-327(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=97124195; PubMed=8969508;
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
 RA Kobayashi Y.,
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
 the *Bacillus subtilis* genome containing the skin element and many
 sporulation genes.",
 RL Microbiology 142:3103-3111(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunze F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 AZEVEDO V., Bertolo M.G., Beesieres P., Bolotin A., Borchert S.,
 BORRIS R., Bourisier L., Brans A., Braun M., Brignell S.C., Bron S.,
 BROUILLET S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 CHOI S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 DENIZOT F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 ENILIAN K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 FRITZ C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
 GHIM S.Y., Glaeser P., Goffeau A., Gollightly B.J., Grandi G.,
 GIMSEPPI G., Guy B.J., Haga K., Hache J., Harwood C.R., Henaut A.,
 HILBERT H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 JORIS B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,
 KOBAYASHI Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
 KURIKA K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 LEE S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
 MEDINA N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 NOONE D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B., Park S.H.,
 PARRIO V., Pohl T.M., Portetelle D., Portollik S., Prescott A.M.,
 PRESCAN B., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RIEGER M., Rivolta C., Rocha B., Roche B., Rose M., Sadale Y.,
 SATO T., Scanlan B., Schleich S., Schroeter R., Scotfione F.,
 SEKIGUCHI J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 SOROKIN A., Taccani E., Takagi T., Takahashi H., Takemaru K.,
 TAKEUCHI M., Tamakoshi A., Tanaka T., Terstra P., Tognoni K.,
 TOSATO V., Uchiyama S., Vandenbol M., Vanier P., Vassarotti A.,
 VIARI A., Wambut R., Wedler E., Wedler H., Weltensger T.,
 WINTERS P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 YOSHIDA K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*

RT "subtilis.",
 RL Nature 390:249-256(1997).
 RN [4]
 RP IDENTIFICATION.
 RX MEDLINE=96084975; PubMed=7489895;
 RA Medigue C., Moszer I., Viari A., Danchin A.,
 RT "Analysis of a *Bacillus subtilis* genome fragment using a co-operative
 RT computer system prototype.",
 RL Gene 165:GC37-GC51(1995)
 CC -1- SIMILARITY: TO B.SUBTILIS YQXC AND T.HYDROXYENTERINE HEMOLYSIN
 CC TLVA.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D32216; BAA06937.1; -
 DR EMBL: D84432; BAA12388.1; -
 DR EMBL: Z99117; CAB14567.1; -
 DR PIR: D69945; D69945.
 DR Subtilist; BG11264; yqam.
 DR InterPro: IPR003593; AAA_Arase.
 DR SMART: SM00382; AAA, 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 313 AA; 3613 MW; 09208795E310D73A CRC64;

Query Match
 Best local similarity 33.1%; Score 127.5; DB 1; Length 313;
 Matches 40; Conservative 19; Mismatches 49; Indels 13; Gaps 6;

QY 98 GEQVKGALYVGFPGTSGFLGAIANOL-KSKRVSTIYLP--EFIRTLKGGFKDGSFE 154
 DB 161 GERQNSIALGQPGSGKTHLTAIMNNILKKSVH--CMYFPVBEQMDLKNF--DNLE 216
 QY 155 KCLHVRKNTIMDDI-----GAEVTPWVADVYIGPLHTRMHELTPFSSNDYSE 209
 DB 217 AKLDAMRKVEVLFIDLPKINGOPRATDW--QVEQIQSVLNYRYNKKLLLSSETLIDE 275

QY 210 L 210
 DB 276 I 276

RESULT 4
 XKDC_BACSU
 ID XKDC_BACSU STANDARD; PRT; 266 AA.
 AC P39782;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phage-like element PSX protein xkdc.
 GN XKDC.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / SO113;
 RX MEDLINE=94364963; PubMed=8083174;
 RA McDowell G.E., Wood H., Devine K.M., McConnell D.J.,
 RT "Genetic control of bacterial suicide: regulation of the induction of
 RT PSX in *Bacillus subtilis*.",
 RL J. Bacteriol. 176:5820-5830(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX Krogan S., O'Reilly M., Nolan N., Devine K.M.,
 RL submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Aevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boutsier L., Brans A., Braun M., Brigelli S.C., Bron S.,
 RA Brouillet S., Bruchet C.V., Caldwell B., Capuano N.J., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings S.D., Emerson P.T.,
 RA Denzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertan K.D., Ertanoglu J., Fabre C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita Y., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giesepp G., Guy B.J., Haga K., Haeck K., Haeck C.R., Henaut A.,
 RA Hilbert H., Holmberg S., Hosono S., Hulio M.P., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klier-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudaga B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Portolillo S., Prescott A.M.,
 RA Presecan E., Pujic P., Ruelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
 RA Sorokin M., Tacconi B., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni A.,
 RA Tosaio V., Uchiyama S., Vandenbol M., Vannier P., Vassartel A.,
 RA Vlati A., Wambutt R., Wedler H., Weitzenecker T.,
 RA Winters P., Wiput A., Yamamoto H., Yamane K., Yaeumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein B., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis".
 RL Nature 390:249-256(1997).
 CC - FUNCTION: MAY FUNCTION AS A TRANSCRIPTIONAL ANTITERMINATOR.
 CC - SIMILARITY: STRONG, TO B.SUBTILIS YQAM.
 CC -
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 CC EMBL; Z34287; CA84044.1; -;
 DR EMBL; Z70177; CA94054.1; -;
 DR EMBL; Z99110; CAB13110.1; -;
 DR PIR; I40411; I40411.
 DR Subtilisin; BG10996; xkdc.
 DR InterPro; IPR003593; AAA_ATPase.
 DR SMART; SM00382; AAA; 1.
 KW ATP-binding; Complete proteome.
 FT NP BIND 124 131 ATP (POTENTIAL).
 FT CONFLICT 239 239 K -> R (IN REF. 1).
 SQ SEQUENCE 266 AA; 30534 MW; 9929C991E9D655AA CRC64;
 Query Match 9.6%; Score 127; DB 1; Length 266;
 Best Local Similarity 31.0%; Pred. No. 0.0019;
 Matches 39; Conservative 25; Mismatches 44; Indels 18; Gaps 7;
 QY 92 CTA--LINGROV-----GLVPGFGTGSFLLGAIANQLKSKVRSTIYL--EFIR 142
 DB 101 CTBYADVADYQIDCKRNSIALGQPSGKTHLTAANEL--MTCYVPIYFPEVGGFT 159
 QY 143 TLKGGFQDGSFEKTLHVRBANILMDI-----GAEVTPWVDEYIGPLHYRMVHEL 197
 DB 160 DLKNDP--ALLEKLNKMKQADYVIFIDDLKPVNKGKRRATDMQLEQYNS--VLAHYRYLNK 216
 QY 198 PTFPSS 203
 DB 217 PILLS 222

RESULT 5
 ID DNA_CAMBE STANDARD; PRT; 440 AA.
 AC GPEJBO;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chromosomal replication initiator protein dnaA.
 GN DNA OR C07001.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jorgensen K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences."
 RL Nature 403:665-668(2000).
 CC - FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
 CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
 CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
 CC (DNA BOX): 5'-TATATC(A)/A/C/A-3'. DNA BINDS TO ATP AND TO
 CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
 CC - SIMILARITY: Belongs to the dnaA family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 CC EMBL; AL139074; CAB72494.1; -;
 DR PIR; D81415; D81415.
 DR HAMAP; MF 00377; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001957; Bac_DnaA.
 DR Pfam; PF00308; bac_dnaA; 1.
 DR PRINTS; PR00051; DNAA.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR00362; DnaA; 1.
 DR PROSITE; PS01008; DNAA; FALSE NEG.
 KW DNA replication; DNA-binding; ATP-binding; Complete proteome.
 FT NP BIND 143 150 ATP (POTENTIAL).
 SQ SEQUENCE 440 AA; 49686 MW; A3FAB90CCED6BD7 CRC64;
 Query Match 9.5%; Score 125.5; DB 1; Length 440;
 Best Local Similarity 26.8%; Pred. No. 0.0048;
 Matches 41; Conservative 30; Mismatches 51; Indels 31; Gaps 7;
 QY 92 CTAITNGEYKGLY---LVPGFGTGSFLLGAIANQLKSKVRSTIYL--DEPIRTLK 145
 DB 123 CKAIANKDKLGLKLYNIPYVGPFGTGLKTHLQAVGN--ASLENGKQVYATSENFINDF 180
 QY 146 GKGQDGSFEKTLHVRBANILMDI-----GAEVTPWVDEYIGPLHYRMVHEL 205
 DB 181 SNKNSGSLDKFHKRYKNCVLLIDYDQFLGKTDKIDEE-----FF---F 221
 QY 206 DYSELSEHH---LAMTPDGEKTKAARILIRVYS 235
 DB 222 IFNEIKNDGQILITSDNPNNMLKG--ITERLKS 253
 RESULT 6

DNAC_BUCAP STANDARD; PRT; 246 AA.

ID DNAC_BUCAP 08KAT9;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE DNA replication protein dnac.

GN DNAC OR BUSG022.

OS Buchnera aphidicola (subsp. Schizaphis graminum).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Buchnera.

NCBI_Taxid=98794;

RP SEQUENCE FROM N.A.

RA MEDLINE=2084549; PubMed=12089438;

RA Tamas I., Klason L., Candace B., Naeelund A.K., Erikson A.-S.,

RT Werngren J.J., Sandstrom J.P., Moran N.A., Andersson S.G.E.,

RL Science 296:2376-2379(2002).

CC -1- FUNCTION: THIS PROTEIN IS REQUIRED FOR CHROMOSOMAL REPLICATION. IT

CC FORMS, IN CONCERT WITH DNAB PROTEIN AND OTHER PREPIMING PROTEINS

CC DNAT, N, N', N' A PREPIMING PROTEIN COMPLEX ON THE SPECIFIC SITE

CC OF THE TEMPLATE DNA RECOGNIZED BY PROTEIN N' (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE DNAC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; AE014077; AAM67594.1; Complete proteome.

KW DNA replication; Primosome; PROBABLY INVOLVED IN THE INTERACTION

FT SITE 69 WITH THE DNAB PROTEIN (BY SIMILARITY).

FT SEQUENCE 246 AA; 28497 MW; BF7E2A9BC2ADB5D CRC64;

QY Query Match 9.2%; Score 122; DB 1; Length 246;

QY Best Local Similarity 23.6%; Pred. No. 0.0046;

QY Matches 48; Conservative 39; Mismatches 86; Indels 30; Gaps 7;

QY 48 DEERFAELITSH---WQRTLNANKDIOYNN-----HRRLOYAMADICRA 94

DB 35 EGRLLSESLREKMKQVLRSGIRLNNCSFENYKIEHDQRTKLAKKAYAE 94

QY 95 ITNGEVOVGLYLPPTGKSFILGAIANQKSKVRSSTIYLPFRITLKGFGKDS-- 152

DB 95 F--NENIASPIFGSGRTGKHLASAIQNLILHGSILIVYADLMSNKKGFSGTSNI 152

QY 153 -FEKKLHVRVREANILMDDIGAEVTPWVDEVIPLHRYMVELPTFPSSNFDYSELE 211

DB 153 TEENLLHNSVLDLMDIDEGMTESRYEK-VIIINOIVDRSSSKSTGMLSLDRGKM 211

QY 212 HHLAMTRDGEKTKARITERRV 234

DB 212 NL-----GS-----RVIDRMR 223

RESULT 7

DNAC_BUCAI STANDARD; PRT; 246 AA.

ID DNAC_BUCAI P57134;

AC P57134;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE DNA replication protein dnac.

GN DNAC OR BU021.

OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum

OS symbiotic bacterium).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Buchnera.

NCBI_Taxid=118099;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Tokyo 1998;

RX MEDLINE=20445173; PubMed=10993077;

RA Shigenobu S., Matnabe H., Hattori M., Sakaki Y., Ishikawa H.;

RT "Genome sequence of the endocellular bacterial symbiont of aphids

RT Buchnera sp. APS."

RL Nature 407:81-86(2000).

CC -1- FUNCTION: THIS PROTEIN IS REQUIRED FOR CHROMOSOMAL REPLICATION. IT

CC FORMS, IN CONCERT WITH DNAB PROTEIN AND OTHER PREPIMING PROTEINS

CC DNAT, N, N', N' A PREPIMING PROTEIN COMPLEX ON THE SPECIFIC SITE

CC OF THE TEMPLATE DNA RECOGNIZED BY PROTEIN N' (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE DNAC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

DR EMBL; AP001118; BAB12748.1; Complete proteome.

KW DNA replication; Primosome; PROBABLY INVOLVED IN THE INTERACTION

FT SITE 69 WITH THE DNAB PROTEIN (BY SIMILARITY).

FT SEQUENCE 246 AA; 28444 MW; C2B56A3BCD1A68F CRC64;

QY Query Match 8.7%; Score 115.5; DB 1; Length 246;

QY Best Local Similarity 22.5%; Pred. No. 0.016;

QY Matches 49; Conservative 40; Mismatches 88; Indels 41; Gaps 8;

QY 44 KRYDEE-----FEAELITSH---WQRTLNANKDIOYNN-----HR 79

DB 20 KEKFDNEDILLAMNOEGRLSESIIRKMKQVLRSGIRLNNCSFENYKIEHD 79

QY 80 DLDVYAMADDTCTATNGEVOVGLYLPPTGKSFILGAIANQKSKVRSSTIYLPB 139

DB 80 GGRKVLKSKRYAER--NENIASPIFGSGRTGKHLASAIQNLILHGSILIVYAD 137

QY 140 FRTLLKGFGKDS--FEKKLHVRVREANILMDDIGAEVTPWVDEVIPLHRYMVE 196

DB 138 LMSNKKGTFSGTSNTEENLHLDLSSVDLMDIDEGMTESRYEK-VIIINOIVDRSSSK 196

QY 197 LPTFPSSNFDYSELEHHLAMTRDGEKTKARITERRV 234

DB 197 RSTGMLSNLDHGMKSL-----GS-----RVIDRMR 223

RESULT 8

DNAA_AQUAE STANDARD; PRT; 399 AA.

ID DNAA_AQUAE O66659;

AC O66659;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chromosomal replication initiator protein dnaa.

GN DNAA OR AQ_322.

OS Aquifex aeolicus.

OC Bacteria; Aquificales; Aquificaceae; Aquifex.

OC NCBI_Taxid=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.B., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swenson R.V.,

RT "The complete genome of the hyperthermophilic bacterium Aquifex

RT aeolicus."

RL Nature 392:353-358(1998).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 76-399.
 RX MEDLINE=22220049; PubMed=12234917;
 RA Erzberger J.P., Pirruccello M.M., Berger J.M.;
 RT "The structure of bacterial DnaA: implications for general mechanisms
 underlying DNA replication initiation".
 RL EMBO J. 21:4763-4773(2002).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
 OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
 CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
 CC (DNAA BOX): 5'-TTATC(C/A)A(C/A)-3'. DNAA BINDS TO ATP AND TO
 CC ACTIDIC PHOSPHOLIPIDS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the dnaA family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AB006683; AAC06612.1; -
 DR PIR: G70328; G70328.
 DR PDB: 1LBQ; 25-SEP-02.
 DR HAMAP: MF_00377; -1.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001957; Bac DnaA.
 DR Pfam: PF00308; bac_dnaa; 1.
 DR PRINTS: PR00051; DNAA.
 DR SMART: SM00382; AAA; 1.
 DR TIGRFAMs: TIGR00362; DnaA; 1.
 DR PROSITE: PS01008; DNAA; 1.
 KM DNA replication; DNA-binding; ATP-binding; Complete proteome;
 KM 3D-structure.
 FT NP_BIND 119 126 ATP (POTENTIAL).
 SQ SEQUENCE 399 AA; 46840 MW; C641CB199AA79906 CRC64;

Query Match 8.7%; Score 115; DB 1; Length 399;
 Best Local Similarity 25.2%; Pred. No. 0.031;
 Matches 33; Conservative 31; Mismatches 53; Indels 14; Gaps 4;

QY 44 KIRYDEERFAELITSHMORDTLNAK-LKDIYNNRDLVAMAAADDICTAITN-GEQ 100
 DB 63 EYKEKEKKKVEI-----KDFLNPKYLTENFVIGGNRL---AYEVKALENLGSL 111
 QY 101 VKGLVLYPGPGTGSFTLGAIANQLKSKVRSITIVPEIRTLKGFRCGSEKKLHRY 160
 DB 112 YNPIFYGVGTGKTHLQAAGNAKRGYRVYSSADPQAQAVEHLKKGTTNEPRMY 171
 QY 161 REANIILADDI 171
 DB 172 KSYVDLLLDV 182

RESULT 9
 ID 1STB_PSEAE STANDARD; PRT; 265 AA.
 AC P15026;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Insertion sequence IS21 putative ATP-binding protein.
 GN 1STB.
 OS Pseudomonas aeruginosa.
 OC Plasmid R68.45.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89218951; PubMed=2540414;
 RA Reilmann C., Moore R., Little S., Savioz A., Willette N.S., Haas D.;
 CC "Genetic structure, function and regulation of the transposable

RT element IS21.";
 RL Mol. Gen. Genet. 215:416-424(1989).
 RN [2]
 RP REVISION TO 283.
 RA Berger B.;
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE 1STB PROTEIN, ONE OF TWO PROTEINS EXPRESSED ONLY
 CC WHEN THERE IS A TANDEM REPEAT OF THE IS21 INSERTION SEQUENCE, IS
 CC NECESSARY FOR THE TRANSPOSITION OF PLASMIDS WITH THAT TANDEM
 CC REPEAT.
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING
 CC PROTEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X14793; CAA32899.2; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR002611; 1stb_ATPbind.
 DR Pfam: PF01695; 1stb; 1.
 DR SMART: SM00382; AAA; 1.
 KM Plasmid; Transposable element; ATP-binding.
 FT NP_BIND 104 111 ATP (POTENTIAL).
 SQ SEQUENCE 265 AA; 30528 MW; 904C7ADC5E3A12A9 CRC64;

Query Match 8.6%; Score 114; DB 1; Length 265;
 Best Local Similarity 26.1%; Pred. No. 0.023;
 Matches 42; Conservative 36; Mismatches 53; Indels 30; Gaps 8;

QY 99 EYKGLVLYPGPGTGSFTLGAIANQLKSKVRSITIVP-BFIRTLKGFRCGSEKK 156
 DB 95 ERSSENVILIGPGGVGKTHL-AIALGVKAVDAGHRYLFMPLRLITLTKAKQENLEKQ 152
 QY 157 LHRVRANIMLDDIGAEVTPWVRDEVIGPLHRYMVA---ELTFPSSN---FDYSE 209
 DB 153 LQQLSYARVLLIDEG--YLPNNHEA--SLFFELNRRYKKSITILTSNGFADWG 206
 QY 210 L-EHILAMTRDGEKTKAARIIRKYSISTYFVLSGENFR 248
 DB 207 MRGDHVL-----ATAILDLRHHSITLINIKGESYR 236

RESULT 10
 ID DNAA_ACHLA STANDARD; PRT; 445 AA.
 AC O9KH08;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chromosomal replication initiator protein dnaA.
 GN DNAA.
 OS Acholeplasma laidlawii.
 OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
 OC Acholeplasmataceae; Acholeplasma.
 OC NCBI_TaxID=2148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PG-8B;
 RX MEDLINE=20242229; PubMed=10779957;
 RA Taganov K.D., Guenchin A.E., Akopian T.A., Oparina N.Y.,
 RA Abgarycheva N.Y., Govorun V.M.;
 RT "Analysis of genes, coding for DNA gyrase from the mycoplasma
 RT Acholeplasma laidlawii PG-8B".
 RL Mol. Biol. (Mosk) 34:292-299(2000).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
 CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
 CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
 CC (DNAA BOX): 5'-TTATC(C/A)A(C/A)-3'. DNAA BINDS TO ATP AND TO

CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).

CC -1- SIMILARITY: Belongs to the dnaa family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF248639; AAF75986.1; -

CC DR HAMAP; MF 00377; -; 1.

CC DR InterPro; IPR003593; AAA_ATPase.

CC DR InterPro; IPR001957; Bac_Dnaa.

CC DR Pfam; PF00308; Bac_dnaa; 1.

CC DR PRINTS; PR00051; Dnaa.

CC DR SMART; SM00382; AAA; 1.

CC DR TIGRFAm; TIGR00362; Dnaa; 1.

CC DR PROSITE; PS01008; Dnaa; 1.

CC DR KEGG; K01008; Dnaa; 1.

CC KM NP_BIND

CC FT NP_BIND 144 151 ATP (POTENTIAL).

CC SQ SEQUENCE 445 AA; 51046 MW; A6045EE53F8C9DB3 CRC64;

Query Match 8.4%; Score 111.5; DB 1; Length 445;

Best Local Similarity 21.8%; Pred. No. 0.069;

Matches 37; Conservative 34; Mismatches 74; Indels 25; Gaps 7;

QY 19 EVKGVPELYVDN-NRIKIRYLOCPCKIKY-DEERFEBELITSHMQDTLNKLDIY-75

DB 50 FLKRNKLYIAIKINELATYSSTPVRLKFSQEEVLEEVADRKLTIDYRQGNLSTYT 109

QY 76 -----MHRRDLVDYMAADDICTAITNGEYKGLYLPFGTGSFLIGAIANOLKSK 128

DB 110 PSFVGVKSMFAPRAKMAKVAHPGAVAN-----PFYIFEDVGLGKTHLMQAGNYILDN 164

QY 129 KYASTIIV-----LPEFIRTL-KGSGFGSGFEKTLHRYEANIIMLDDI 171

DB 165 DVKRLIYKADNFIEDFVLSRNKNTKEFPNAK--YQDIDVILVDI 211

RESULT 11

ISTR BURCE STANDARD; PRT; 229 AA.

AC P55923;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Insertion sequence IS408 putative ATP-binding protein (ORF2).

OS Burkholderia cepacia (Pseudomonas cepacia).

OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

OC Burkholderiaceae; Burkholderia.

OX NCBI_TaxID=292;

RM [1]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 17616 / 249;

RX MEDLINE=94302134; PubMed=7518087;

RA Byrne A.M., Leesele T.G.;

RT "Characteristics of IS401, a new member of the IS3 family implicated

RT in plasmid rearrangements in Pseudomonas cepacia.";

RL Plasmid 31.138-147(1994).

CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING

CC PROTEIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; L09108; -; NOT_ANNOTATED_CDS.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR002611; IserB_ATPbind.

DR Pfam; PF01695; IserB; 1.

DR SMART; SM00382; AAA; 1.

KW Transposable element; ATP-binding.

FT NP_BIND 105 112 ATP (POTENTIAL).

SQ SEQUENCE 229 AA; 25929 MW; ACS5D619C729163E1 CRC64;

Query Match 8.2%; Score 109.5; DB 1; Length 229;

Best Local Similarity 27.0%; Pred. No. 0.045;

Matches 50; Conservative 32; Mismatches 80; Indels 23; Gaps 8;

QY 59 SHMQDITLNAIKK-----DIYMHRRDLVDYMAADDICTAITNGEYKGLYLYG 108

DB 51 TRRLRLRAALTKRQPCVDEIHYRQTRGLDQRYVATLGG--CDWVRAGQ--NLILTG 105

QY 109 PFGTGSFLIGAIANOLSKKVRSTIIVLPEFIRTLKGSGFEKTLHRYEANIIML 168

DB 106 PFGAGKTWLACAFGQACQSGSVFVYVARLFEELKIHGSGSFTRLAQIAKIDVIL 165

QY 169 DDIQAEVYTPWVRDVIYGLHYRWVHELPTFFSSNPDQSELEHIAWTRDDEKTKAR 228

DB 166 DDMGQDLDQQAARNDL-EVLDDR-VGRSTVITSOL--PLEHWAWLQD--PTLADA 217

QY 229 IIEYV 233

DB 218 IIDRL 222

RESULT 12

DNA PASMU STANDARD; PRT; 451 AA.

AC 09CQ4;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chromosomal replication initiator protein dnaa.

GN DNA OR PM1161.

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Pasteurella.

OX NCBI_TaxID=747;

RM [1]

RP SEQUENCE FROM N.A.

RX STRAIN=Pm70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.U., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;

RT "Complete genomic sequence of Pasteurella multocida Pm70.";

RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION

CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION

CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS

CC (DNAA BOX): 5'-TTATC(C/A)A(C/A)-3'. DNAA BINDS TO ATP AND TO

CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).

CC -----

CC -1- SIMILARITY: Belongs to the dnaa family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; AE006156; AAK03245.1; -

CC DR HAMAP; MF 00377; -; 1.

CC DR InterPro; IPR003593; AAA_ATPase.

CC DR InterPro; IPR001957; Bac_Dnaa.

CC DR Pfam; PF00308; Bac_dnaa; 1.

CC DR PRINTS; PR00051; Dnaa.

CC DR SMART; SM00382; AAA; 1.

CC DR TIGRFAm; TIGR00362; Dnaa; 1.

CC DR PROSITE; PS01008; Dnaa; 1.

KM DNA replication; DNA-binding; ATP-binding; Complete proteome.
 NP BIND 156 ATP (POTENTIAL).
 SQ SEQUENCE 451 AA; 51378 MW; 76A0ACD237F7FF2 CRC64;

Query Match 8.1%; Score 107; DB 1; Length 451;
 Best Local Similarity 21.3%; Pred. No. 0.16;
 Matches 52; Conservative 47; Mismatches 93; Indels 52; Gaps 12;

QY 19 FVKGHVELYV-----DNNRIKIRYLCPCKKYDEERFEALITSHHMDTL 67
 DB 50 FIKQKDESYLAQILTKVAQELSGNAELV---QKVGKPEPKQAQSAALTTNNKEBK 106
 QY 68 NAKLKDIYMNDRDL--VMAAADICTAINT-----GEQVKG-LYLYGPFQKSPIL 118
 DB 107 PQVIRSYLNPKHVFNPEVKGSKNQALARAQVADNPGEPSNPLFLYGSTGLGKTHL 166
 QY 119 GAIANOLSKKVRSTIYL--PEFIRTLKGFQKGSFE--KKLRVREANILMDIT--- 171
 DB 167 HAIGNGILSRNTARVLYIHANNFMQMVANVRDNKDEKCFYRSIDA--LVDDIQPF 224
 QY 172 GAEVTPWVNDVIGPLIHYRVHLELPPFSNFPDYSLEHHLAMTRDGEKTKAARIIE 231
 DB 225 AEKKTQ-----EEFPIHFNLPDTG---RQITLTSDRTPK-EIEKLEB 264
 QY 232 RVKS 235
 DB 265 RLKS 268

RESULT 13
 ID 1STB_BACST STANDARD; PRT; 251 AA.
 AC Q45619;
 DT 01-NOV-1997 (Rel. 35, Last Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Insertion sequence IS5376 putative ATP-binding protein.
 OS Bacillus stearothermophilus.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 NC NCBL_TaxID=1422;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CU21;
 RC MEDLINE=93181468; PubMed=8382825;
 RA Xu K., He Z.-Q., Mao Y.-M., Shen R.-Q., Sheng Z.-J.;
 RT "On two transposable elements from Bacillus stearothermophilus.";
 RL Plasmid 29.1-9(1993).
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X67861, CAA48046.1, -.
 DR PTR; S23889;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR002611; IctB_ATPbind.
 DR Pfam: PF01695; IctB; 1.
 DR SMART: SM00382; AAA; 1.
 KM Transposable element; ATP-binding.
 NP BIND 105 112
 SQ SEQUENCE 251 AA; 29286 MW; FA8FP4F55B2FB527 CRC64;

Query Match 8.0%; Score 106.5; DB 1; Length 251;
 Best Local Similarity 22.1%; Pred. No. 0.089;
 Matches 49; Conservative 40; Mismatches 98; Indels 35; Gaps 8;

QY 45 IKYDERF---EALITSHHMDTLNAKLDIYMNDRDLVAM--AAD--ICTAINT 97

DB 33 ISVSEFLRLLEAEIVKQARSIGTLIKLSKLPYRKTIIDFDTAQSVDERRIRELTL 92
 QY 98 G--EQVKGLYLNPFGFGKSPFIIAGAIANOLSKKVRSTIYLPFIRTLKGFQDSFEK 155
 DB 93 SFIDRENKILFLDPGGIKTHLAIISGWEIAGYKFTTADLVQLRAAQEGLEK 152
 QY 156 KLRVREANILMDIDGAEVTPWVNDVIGPLIHY-----RMVHELPPFSNFPDYSR 209
 DB 153 KLRVFPKPTVLIIDEMQYIKLDP-----NSAHYFQVIARVYHAPLIILSNKSPGE 204
 QY 210 LEHHLAMTRDGE--EKTVAARIERYKSLSTPYFSGENFR 248
 DB 205 W-----GRIYGSVLATAMDLRLHSHIIFNLKGESEYR 237

RESULT 14
 ID 1STB_PSEFL STANDARD; PRT; 231 AA.
 AC Q51762;
 DT 01-NOV-1997 (Rel. 35, Last Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Insertion sequence IS1162 putative ATP-binding protein.
 OS Pseudomonas fluorescens.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
 NC NCBL_TaxID=294;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ST;
 RC MEDLINE=95212933; PubMed=7698671;
 RA Solinas F., Marconi A.M., Ruzzi M., Zennaro B.;
 RT "Characterization and sequence of a novel insertion sequence, IS1162, from Pseudomonas fluorescens.";
 RL Gene 155:77-82(1995).
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING PROTEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X79443; CAA55960.1, -.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR002611; IctB_ATPbind.
 DR Pfam: PF01695; IctB; 1.
 DR SMART: SM00382; AAA; 1.
 KM Transposable element; ATP-binding.
 NP BIND 107 114
 SQ SEQUENCE 231 AA; 25997 MW; 341592AEA7DD794A CRC64;

Query Match 7.9%; Score 105.5; DB 1; Length 231;
 Best Local Similarity 21.4%; Pred. No. 0.097;
 Matches 33; Conservative 33; Mismatches 61; Indels 27; Gaps 3;

QY 35 KIRYLQCPCKIKYDERFEALITSHHMDTLNAKLDIYMNDRDLVAMAADICTA 94
 DB 55 RLRLKQKQKTK-----ASACLEDIDRRRRGRITLDERLTA-----T 90
 QY 95 ITNGEYK---GLYVPGFGTGSFIIAGAIANOLSKKVRSTIYLPFIRTLKGFQK 151
 DB 91 LNSGDIRQRHNLILGPGVGTWPCALGHWACQGSALVLRTPRLLEQRIAHG 150
 QY 152 SPEKILRVREANILMDIDGAEVTPWVNDVIT 185
 DB 151 SFGRITQQLAKVDVLTLDWGIALAEENARHDL 184

Search completed: December 15, 2003, 15:14:19
 Job time : 8.90922 secs

```

RESULT 15
YDAB_ECOLI          STANDARD;          PRT;          248 AA.
ID   YDAB_ECOLI
AC   P75746
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DE   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein ydab.
GN   YDAB OR B1360
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
[1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=9742617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
[2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97251357; PubMed=9097039;
RA   Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA   Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA   Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA   Nakade S., Nakamura Y., Nashimoto H., Nishio Y., Oshima T., Saito N.,
RA   Sasaki G., Seki Y., Sivasubramanian S., Tagami H., Takeda J.,
RA   Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RT   "A 570-Kb DNA sequence of the Escherichia coli K-12 genome
RT   corresponding to the 28.0-40.1 min region on the linkage map.";
RL   DNA Res. 3:363-377(1996).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; AE000233; AAC74442.1; -
DR   EMBL; D90774; BAA14958.1; -
DR   PIR; C64886; C64886.
DR   EcoGene; BG13364; Ydab.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR002611; IctB_ATPbind.
DR   Pfam; PF01695; IctB_1.
DR   SMART; SM00382; AAA; 1.
KW   Hypothetical protein; Coiled coil; Complete proteome.
FT   DOMAIN          33..57       COILED COIL (POTENTIAL).
SQ   SEQUENCE        248 AA; 28128 MW; 3501DCB1E8512E01 CRC64;

Query Match          7.3%; Score 105; DB 1; Length 248;
Best Local Similarity 26.7%; Pred. NO. 0.12;
Matches 43; Conservative 30; Mismatches 66; Indels 22; Gaps 7;

QY 77 NHEDRLDVMA---ADDICAIITNGEYVKGILYIPRGTKSPILGAINQLSKKYRST 133
DB 80 NEGQRYALSOAKSIADBLMTGCTN-----FAPSGKPGTGKNDLAALIGNRLLKDGQTVI 133
QY 134 IIVLPFIRITLKGFKDG-SFEKKLHVRBANILMLDDIGAEVTPWVRDEVIGPLHYR 192
DB 134 VVTVADVMSALHLSYDDGSGEKKFRLCEVDLVLDEIGIQRET---KNEQV--VLH-Q 187
QY 193 MVEHELPTFFSSNDYSELEHHLAMTRDGEKTKAARIERY 233
DB 188 IYDRRTASMRSSVGMILTNLNTEAMKTLIGE-----RIMDRM 222
  
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 28.8858 Seconds
(without alignments)
2233.384 Million cell updates/sec

Title: US-09-689-952-18
Perfect score: 1328
Sequence: 1 YKDQKXHDGKHFADCPNFV.....ERVKSLSTPYPLSGNPFANN 250

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL, 23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1328	100.0	306	16	Q99TH8
2	1320	99.4	306	16	Q8NW67
3	1189	89.5	306	16	Q8CNV2
4	531.5	40.0	311	16	Q9K863
5	523	39.4	308	16	Q8CX80
6	475.5	35.8	307	16	Q92BF4
7	472.5	35.6	307	16	Q8Y6X1
8	397.5	29.9	293	16	Q8B2J5
9	397.5	29.9	300	16	Q9A1D1
10	392.5	29.6	299	16	Q8CWM7
11	381	28.7	298	16	Q8CWP7
12	365	27.5	300	16	Q8S1T8
13	361	27.2	300	16	Q8DY72
14	342	25.8	293	16	Q97PC7
15	342	25.8	293	16	Q9CHH9
16	205.5	15.5	74	2	Q05653

17	194	14.6	343	16	Q8EWK0	Q8ewk0 mycoplasma
18	173	13.0	299	2	Q48991	Q48991 mycoplasma
19	157	11.8	297	2	Q48R03	Q48r03 mycoplasma
20	153	11.5	259	9	Q96022	Q96022 bacterioph
21	151	11.4	259	9	Q9B0F8	Q9b0f8 staphylococ
22	145	10.9	261	16	Q932A3	Q932a3 staphylococ
23	145	10.9	282	16	Q97HS5	Q97hs5 clostridium
24	140	10.5	257	9	Q9WBR8	Q9wbr8 staphylococ
25	140	10.5	316	16	Q9P0I4	Q9p0i4 ureaplasma
26	138.5	10.4	235	16	Q67056	Q67056 aquifex ae
27	129	9.7	285	16	Q8E256	Q8e256 leptospira
28	127	9.6	249	2	Q93SD5	Q93sd5 escherichia
29	120.5	9.1	316	2	P95446	P95446 pseudomonas
30	114	8.6	457	16	Q8XPG2	Q8xpg2 clostridium
31	113.5	8.5	195	17	Q8PTM8	Q8ptm8 methanobact
32	113.5	8.5	257	17	Q8TWM8	Q8twm8 methanobact
33	113.5	8.5	257	17	Q8TH29	Q8th29 methanobact
34	113	8.5	327	16	Q8R6U1	Q8r6u1 thermomanae
35	113	8.5	796	17	Q8TZV0	Q8tzv0 pyrococcus
36	112.5	8.5	261	2	Q8G101	Q8g101 pseudomonas
37	112	8.4	287	16	Q926A4	Q926a4 listeria in
38	112	8.4	291	16	Q9C1B5	Q9c1b5 lactococcus
39	112	8.4	294	9	Q8L7N1	Q8l7n1 lactococcus
40	112	8.4	294	9	Q9AZP7	Q9azp7 bacterioph
41	112	8.4	520	17	Q26312	Q26312 methanobact
42	111.5	8.4	296	9	Q38098	Q38098 bacterioph
43	111.5	8.4	475	10	Q9LPI1	Q9lpi1 arabidopsis
44	111	8.4	262	2	Q93P71	Q93p71 microscilla
45	111	8.4	350	16	Q921Y8	Q921y8 rickettsia

ALIGNMENTS

RESULT 1
ID Q99TH8 PRELIMINARY: PRT: 306 AA.

AC Q99TH8: 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Primosomal protein.
GN DNAI OR SAV1684 OR SA1507.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_Taxid=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECTRUS-S. aureus (strain Mu50), and S. aureus (strain N315).
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Itoh T., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshoyama A.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Kanihara M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayaishi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RT Lancet 357:1225-1240(2001).
DR EMBL; AP003363; BAB57846.1; -;
DR EMBL; AP003134; BAB42774.1; -;
KW Complete proteome.
SQ SEQUENCE 306 AA; 35635 MW; 7F3440E89643505E CRC64;

Query Match 100.0%; Score 1328; DB 16; Length 306;
Best Local Similarity 100.0%; Pred. No. 7.3e-107;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 YKDQKXHDGKHFADCPNFVKGVPELVYDNNRIKIRYLQCCCKIKYDERFEALITSH 60
DB 57 YKDQKXHDGKHFADCPNFVKGVPELVYDNNRIKIRYLQCCCKIKYDERFEALITSH 116

QY 61 HMORDTLNKLDKIYNNHRLDVAMAADDICTAINTGEOVKGLYLGPFGTGSFTLGA 120
 DB 117 HMORDTLNKLDKIYNNHRLDVAMAADDICTAINTGEOVKGLYLGPFGTGSFTLGA 176
 QY 121 INANQSKKVRSTIYLPFRITLKGFGKGSFEKKLHVRREANIIMLDDIGAEVTPMW 180
 DB 177 INANQSKKVRSTIYLPFRITLKGFGKGSFEKKLHVRREANIIMLDDIGAEVTPMW 236
 QY 181 RDEVIGPLHYRMVHLPFPSSNFDYSELEHLLAMTRDGEKTKAARIIRKXSLSTPY 240
 DB 237 RDEVIGPLHYRMVHLPFPSSNFDYSELEHLLAMTRDGEKTKAARIIRKXSLSTPY 296
 QY 241 FLSGENFRNN 250
 DB 297 FLSGENFRNN 306

RESULT 2

Q8NM67 PRELIMINARY; PRT; 306 AA.
 ID Q8NM67
 AC Q8NM67; 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE Primosomal protein.
 GN DNAI OR MM1627.
 OS Staphylococcus aureus (strain MW2).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NCBI TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwano N., Asano K., Naito T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA".
 RL Lancet 359:1819-1827(2002).
 DR EMBL; AP004627; BAB95492.1; -.
 KW Complete proteome.
 SQ SEQUENCE 306 AA; 35626 MW; C0FCA3752D934B4 CRC64;

Query Match 99.4%; Score 1320; DB 16; Length 306;
 Best Local Similarity 99.6%; Pred. No. 3.6e-106;
 Matches 249; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKDOQKHVDGKRFADCPNFKGVHPELYVDNNRIKIRYLOCPCKIKYDERFEALITSH 60
 DB 57 YKDOQKHVDGKRFADCPNFKGVHPELYVDNNRIKIRYLOCPCKIKYDERFEALITSH 116
 QY 61 HMORDTLNKLDKIYNNHRLDVAMAADDICTAINTGEOVKGLYLGPFGTGSFTLGA 120
 DB 117 HMORDTLNKLDKIYNNHRLDVAMAADDICTAINTGEOVKGLYLGPFGTGSFTLGA 176
 QY 121 INANQSKKVRSTIYLPFRITLKGFGKGSFEKKLHVRREANIIMLDDIGAEVTPMW 180
 DB 177 INANQSKKVRSTIYLPFRITLKGFGKGSFEKKLHVRREANIIMLDDIGAEVTPMW 236
 QY 181 RDEVIGPLHYRMVHLPFPSSNFDYSELEHLLAMTRDGEKTKAARIIRKXSLSTPY 240
 DB 237 RDEVIGPLHYRMVHLPFPSSNFDYSELEHLLAMTRDGEKTKAARIIRKXSLSTPY 296
 QY 241 FLSGENFRNN 250
 DB 297 FLSGENFRNN 306

RESULT 3
 Q8CNY2 PRELIMINARY; PRT; 306 AA.
 ID Q8CNY2
 AC Q8CNY2; 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Primosomal protein.
 GN SE1358.
 OS Staphylococcus epidermidis.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NCBI TaxID=1282;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 12228;
 RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
 RA Chen Z., Wen Y.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016748; AA004957.1; -.
 KW Complete proteome.
 SQ SEQUENCE 306 AA; 35736 MW; EDD9F7B52DBFCA16 CRC64;

Query Match 89.5%; Score 1189; DB 16; Length 306;
 Best Local Similarity 87.2%; Pred. No. 7.9e-95;
 Matches 218; Conservative 23; Mismatches 9; Indels 0; Gaps 0;

QY 1 YKDOQKHVDGKRFADCPNFKGVHPELYVDNNRIKIRYLOCPCKIKYDERFEALITSH 60
 DB 57 YKDOQKHVDGKRFADCPNFKGVHPELYVDNNRIKIRYLOCPCKIKYDERFEALITSH 116
 QY 61 HMORDTLNKLDKIYNNHRLDVAMAADDICTAINTGEOVKGLYLGPFGTGSFTLGA 120
 DB 117 HMORDTLNKLDKIYNNHRLDVAMAADDICTAINTGEOVKGLYLGPFGTGSFTLGA 176
 QY 121 INANQSKKVRSTIYLPFRITLKGFGKGSFEKKLHVRREANIIMLDDIGAEVTPMW 180
 DB 177 INANQSKKVRSTIYLPFRITLKGFGKGSFEKKLHVRREANIIMLDDIGAEVTPMW 236
 QY 181 RDEVIGPLHYRMVHLPFPSSNFDYSELEHLLAMTRDGEKTKAARIIRKXSLSTPY 240
 DB 237 RDEVIGPLHYRMVHLPFPSSNFDYSELEHLLAMTRDGEKTKAARIIRKXSLSTPY 296
 QY 241 FLSGENFRNN 250
 DB 297 YLTKNFRNN 306

RESULT 4

Q9K863 PRELIMINARY; PRT; 311 AA.
 ID Q9K863
 AC Q9K863; 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Primosome component (Helicase loader).
 GN DNAI OR BH3144.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI TaxID=8665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takai H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis".
 RT Nucleic Acids Res. 28:4317-4331(2000).
 RL EMBL; AP001517; BAB06863.1; -.
 DR InterPro; IPR003593; AAA_ATPase.
 DR SMART; SM00382; AAA; 1.
 KW Helicase; Complete proteome.
 SQ SEQUENCE 311 AA; 35822 MW; 6AFB3073431363BB CRC64;

Query Match 40.0%; Score 531.5; DB 16; Length 311;
 Best Local Similarity 41.2%; Pred. No. 6.9e-38;
 Matches 107; Conservative 55; Mismatches 79; Indels 19; Gaps 6;


```

QY 1 YKDOQKHVDGH--KPADCPNFYKGVHVELYVNNRIRKIRYLOCPCKIKYDEERFEALIT 58
DB 59 YQVGEQHC-AHCGLOKCPYLMKGYOPTLYVEDSLSTSPPLKREERREKKSLLR 117
QY 59 SHHMDRTLNKAKLDI-----YNNHRDLVAMADDICTAITNGEQVGYLYGPRG 111
DB 118 SLIYPRLEIAKPDVSEBRSIASHR-ALBFALSAKP-----GEOGMGLYLGKGG 169
QY 112 TQKSPFLGAINOLKSKKVRSTIYLPFRITLKGFQKGSFEKLLHRRREANILMDI 171
DB 170 VGTFTLMGAIAMELKRGIDSTIYVYDFPRELKQSIGDTPQOKIDPVYNAQVLLFDI 229
QY 172 GAEKVTPWVRDEVIGPLHYRMVHEHLPTEPSSNPDYSELEHNLAMT-RDGEKTKARIT 230
DB 230 GAETMTSWRADDVGLYLOTRIMEKLPFTFTSNYDIDELSHLAVNDKSTELLAKRVA 289
QY 231 ERYKSLSTPYLGSNGFRNN 250
DB 290 ERIRHYTVSMVGGQRREH 309

RESULT 5
Q8CB0 PRELIMINARY; PRT; 308 AA.
AC 08CB0;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Primosome component (Hellicase loader).
GN Dnal OR OB2156.
OS Oceanobacillus theyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
ON NCBI_Taxid=182710;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=HTEB31 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RA Genome sequence of Oceanobacillus theyensis isolated from the Itheya
RA Ridge and its unexpected adaptive capabilities to extreme
RA environments."
RA Nucleic Acids Res. 30:3927-3935(2002).
RL EMBL; AP004600; BAC14112.1; -.
DR Hellicase; Complete proteome.
KW Hellicase; Complete proteome.
SQ SEQUENCE 308 AA; 35729 MW; 8D4C2F70E480AE33 CRC64;

Query Match 39.4%; Score 523; DB 16; Length 308;
Best Local Similarity 42.4%; Pred. No. 3, 7e-37;
Matches 106; Conservative 51; Mismatches 91; Indels 2; Gaps 2;

QY 1 YKDOQKHVDGH-KPADCPNFYKGVHVELYVNNRIRKIRYLOCPCKIKYDEERFEALIT 59
DB 60 YKSQSKQCRCKSLGCGQNMIOGYSPYLEADNNDIRLSYKCKHREBQNEQOKLID 119
QY 60 HHMQRTLNKAKLDIYNNHRDLVAMADDICTAITNGEQVGYLYGPRGSKSLTP 119
DB 120 LYMPEKILQARISDVQDEH-RSNALGKVLDFLEAKRQELPKGLVYGSFGVGYLLG 178
QY 120 AINQKSKKVRSTIYLPFRITLKGFQKGSFEKLLHRRREANILMDI GAEVTPW 179
DB 179 AINELKLEYSISLYMPEFVAEIKSSPFQDSFNEKVPFKADILMDKALMQSAN 238
QY 180 VRDEVIGPLHYRMVHEHLPTEPSSNPDYSELEHNLAMTRDGEKTKARITERYKSLSTP 239
DB 239 FRDEVIGSVLYQYRMMEGLPVFITSNYDLDTLEQELSTTRNGVQVAGRIIERIKQVTTD 298
QY 240 YELSGENFRN 249
DB 299 VKLSGPNRS 308

RESULT 6

```

```

Q92BF4 PRELIMINARY; PRT; 307 AA.
AC 092BF4;
ID 092BF4;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Primosome component (Hellicase loader) Dnal.
GN Dnal OR L1N1595.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
ON NCBI_Taxid=1642;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=CLIP 11262 / Serovar 6a;
RX PubMed=11679669;
RA Glaeser P., Frangoul L., Buchrieser C., Rueniock C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Darvar A., Deloux P.,
RA Domann B., Dominguez-Bernal G., Duchaud E., Durant L., Duesugret O.,
RA Ertlan K.-D., Fehli H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krell J., Kuhn M., Kuntz P., Kurapkhat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati A.,
RA Nordiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RA "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596169; CAC96826.1; -.
DR Listeriac; L1N1595; -.
DR InterPro; IPR003593; AAA_Arpase.
DR SMART; SM00382; AAA; 1.
KW Hellicase; Complete proteome.
SQ SEQUENCE 307 AA; 35988 MW; A8517663BA8F123F CRC64;

Query Match 35.8%; Score 475.5; DB 16; Length 307;
Best Local Similarity 42.6%; Pred. No. 4, 8e-33;
Matches 98; Conservative 45; Mismatches 84; Indels 3; Gaps 2;

QY 22 GHVPELYVNNRIRKIRYLOCPCKIKYDEERFEALITSHHMDRTLNKAKLDIYNNHRDR 81
DB 77 GYAPKVLNGBRITVYVYPTKKEIEBKRAVARRIRSLYMRQVADAMIAFYDSESR 136
QY 82 LVVAMADDICTAI--TNGEQVGYLYGPRGSKSLGAINOLKSKKVRSTIYLP 139
DB 137 KALVAYOFLNNYPPQNEERKGLPTGSGSYLGLAKELALGISTITLYLPE 196
QY 140 FRTLKGFQKGSFEKLLHRRREANILMDI GAEVTPWVRDEVIGPLHYRMVHEHLP 199
DB 197 FRYEKQKSDSTVGEKIOPAETEVLMDDIGAESMTMTDDEVYALGILQFRMOBELPT 256
QY 200 PSSNPDYSELEHNLAMTRDG-EKTKARITERYKSLSTPYLGSNGFRN 248
DB 257 PSSNPNMDLENHLMFAQNGTEKELKARIMERYALISKEYVLEBKNR 306

RESULT 7
Q8Y6X1 PRELIMINARY; PRT; 307 AA.
AC 08Y6X1;
ID 08Y6X1;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Primosome component (Hellicase loader) Dnal.
GN Dnal OR LMO1560.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
ON NCBI_Taxid=1639;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=BGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaeser P., Frangoul L., Buchrieser C., Rueniock C., Amend A.,

```

RA Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T.,
 RA Charbit A., Checouani F., Couve E., de Darvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durand L., Duseigneur O.,
 RA Ercian K.-D., Fehli H., Garcia-del Portillo F., Garrido P.,
 RA Gaultier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kuypak G.,
 RA Madueno E., Matouman A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
 RA Rammel B., Rose M., Schlueder T., Symes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Coesart P.,
 RT "Comparative genomics of *Listeria species*,"
 RL Science 294:849-852(2001).
 DR EMBL: AL591979; CAC99638.1; -
 DR L1actidist; LMO01560; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR SMART:SM00382; AAA_1.
 KW Helicase; Complete proteome.
 SQ SEQUENCE 307 AA; 35970 MW; CF2FD9350CEB0F48 CRC64;

Query Match 35.6%; Score 472.5; DB 16; Length 307;
 Best Local Similarity 42.1%; Pred. No. 8,7e-33;
 Matches 98; Conservative 47; Mismatches 79; Indels 9; Gaps 3;

QY 22 GHVPELVYDNNRIRKIRYLQCPCKIKYDEERFEALITSHMORDTLNAKLDIYMRDR 81
 DB 77 GYAPKVLNAGEFIVTYTYPTKEKIEEDKRAVERRIRSLYMPQVVDANLADFTDESR 136
 QY 82 LDVAMADDICTAITN-----GEQVKGILYXPGPTGKSLTILANOLSKYKSTIY 136
 DB 137 ---QLALVEYQFLNANNPPKSGRVRKGLFTHGSGTGKSYLLAALAEKLGISTLLV 193
 QY 137 LPEFIRTLKGFPGKSGFEKTLHVRANILMLDIGAEEVTPWVRDEVIGPLHYRWVHE 196
 DB 194 LPEFMRVKSISDNTYGEKIQPAKREVLMLDIGESTANTATRDVLAIIQFRNGE 253
 QY 197 LPTFFSSNFDYSELEHNLAMTRDGEKTKAARIIEYKSLSTPYLSENGFR 248
 DB 254 LPTFFSSNFMDDJENHLMFAONGTEKELKARRIMERVRLSKVNLGNRR 306

RESULT 8
 Q8P2J5 PRELIMINARY; PRT; 293 AA.
 ID Q8P2J5
 AC Q8P2J5;
 DT 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 GN Putative primosome component (helicase loader).
 GN SPYM18_0332.
 OS Streptococcus pyogenes (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCB1_TaxID=186103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS8232 / serotype M18;
 RX MEDLINE=21977593; PubMed=11917108;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Studevant D.B., Ricklets S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
 RT "Genome sequence and comparative microarray analysis of serotype M18
 RT group A Streptococcus strains associated with acute rheumatic fever
 RT outbreaks,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 DR EMBL: AE009978; AAL97087.1; -
 KW Helicase; Complete proteome; Hypothetical protein.
 SQ SEQUENCE 293 AA; 33357 MW; 0E87AD94BDB59986 CRC64;

Query Match 29.9%; Score 397.5; DB 16; Length 293;
 Best Local Similarity 35.9%; Pred. No. 2.5e-26;
 Matches 84; Conservative 56; Mismatches 91; Indels 3; Gaps 3;

QY 17 PNFV-KGHVPELVYDNNRIRKIRYLQCPCKIKYDEERFEALITSHMORDTLNAKLDIY 75
 DB 61 PSYIANGYQPIILAMNGYADVSTLEKELVEAKQAISERIQVLSPSYRHHISDID 120
 QY 76 MNRDLDVAMADDICTAITNGEQVKGILYGPFGTGSFLIGAIANOLSKK-VRSSTI 134
 DB 121 VNNAASMEAFSAIIDEVEQYPSAEQ-KGLYLGDMGIGSYLLAAMAHSEKGVSTTL 179
 QY 135 IYLPFIRTLKGFPGKSGFEKTLHVRANILMLDIGAEEVTPWVRDEVIGPLHYRWV 194
 DB 180 LHPSPALDVKAISGVSVEIDAVKNVPLILDDIGABQATSWRDEVLYILOYRML 239
 QY 195 HELPTFFSSNFDYSELEHNLAMTRDGEKTKAARIIEYKSLSTPYLSENGFR 248
 DB 240 EELPTFTFNYSRADLERKMATIKGSDEYVQAKRWVERRYLARBEPHGANRR 293

RESULT 9
 Q9A1D1 PRELIMINARY; PRT; 300 AA.
 ID Q9A1D1
 AC Q9A1D1;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-OCT-2002 (TREMBlrel. 17, Last sequence update)
 DE Putative primosome component (helicase loader).
 GN DNAI OR SPY0340 OR SPYM3_0248.
 OS Streptococcus pyogenes, and
 OS Streptococcus pyogenes (serotype M3).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NC NCB1_TaxID=1314, 198466;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. pyogenes; STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Ferreira J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Setz S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.,
 RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes,"
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S. pyogenes; STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=12122206;
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.;
 RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence,"
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 DR EMBL: AE006498; AAK33392.1; -
 DR EMBL: AE014141; AAM78855.1; -
 KW Hypothetical protein; Helicase; Complete proteome.
 SQ SEQUENCE 300 AA; 34146 MW; 489B27A789338C5A CRC64;

Query Match 29.9%; Score 397.5; DB 16; Length 300;
 Best Local Similarity 35.9%; Pred. No. 2.6e-26;
 Matches 84; Conservative 56; Mismatches 91; Indels 3; Gaps 3;

QY 17 PNFV-KGHVPELVYDNNRIRKIRYLQCPCKIKYDEERFEALITSHMORDTLNAKLDIY 75
 DB 68 PSYIANGYQPIILAMNGYADVSTLEKELVEAKQAISERIQVLSPSYRHHISDID 127
 QY 76 MNRDLDVAMADDICTAITNGEQVKGILYGPFGTGSFLIGAIANOLSKK-VRSSTI 134
 DB 128 VNNAASMEAFSAIIDEVEQYPSAEQ-KGLYLGDMGIGSYLLAAMAHLSKKGISTTL 186
 QY 135 IYLPFIRTLKGFPGKSGFEKTLHVRANILMLDIGAEEVTPWVRDEVIGPLHYRWV 194
 DB 187 LHPSPALDVKAISGVSVEIDAVKNVPLILDDIGABQATSWRDEVLYILOYRML 246

QY 195 HELPFSSNFDYSELEHHLAMTRDGEKTKARITERYKSLTPYLGSENR 248
 DB 247 BELPTFTSYNSADLERKWKATTKGSDETWQAKRVMERVAYLAKREPHLECANRR 300

RESULT 10

08CWP7 PRELIMINARY; PRT; 299 AA.

AC 08CWP7; PRELIMINARY; PRT; 299 AA.
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Putative DNA replication protein, primosome component (Helicase loader)
 GN DNAS OR SMU.1921.
 OS Streptococcus mutans.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN NCBI_Taxid=1309;
 RP SEQUENCE FROM N.A.
 RC STRAIN=UAI59 / ATCC 700610 / Serotype C;
 RX MEDLINE=2295063; PubMed=12397186;
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J., Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y., Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental pathogen."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
 DR EMBL; AB015016; AAN59532.1; -
 KW Helicase: Complete proteome.
 SQ SEQUENCE 299 AA; 34410 MW; 98CAAB56F5524BD1 CRC64;

Query Match 29.6%; Score 392.5; DB 16; Length 299;
 Best Local Similarity 34.5%; Pred. No. 7.1e-26;
 Matches 86; Conservative 54; Mismatches 104; Indels 5; Gaps 3;

QY 1 YKDOQKHVGHKADCPNFKVHPELYVDNRIKIRYLQCPCKIYDEBRFAELITSH 60
 DB 55 YRLERRFRFNRH--DQAYIAKGYQPLVNMGVADVAAYKTKELIAAKTQAISDINNV 111
 QY 61 HMQORDTLANKKDIYVNMHRLDVMAADICTAINGEQVGLVYGPFGTKSFTLGA 120
 DB 112 SLPRSKYKNSFDINDVDKRLVFKRVADPVQYRPED-KGLVYIGDMGKSTYUWA 170
 QY 121 IANQLSKK-VRSITIIYLPFIFTLKGFQSGFPEKRLHVRANILMLDDIGAEVTPW 179
 DB 171 MAHELSEQGAATTLTHPSPFTIDVKNALINTGVKKEIDAVKTADILLIDIGAEQSTW 230
 QY 180 VRDEVIGPLIHYMVELPFSSNNDYSELEHHLAMTRDGEKTKARITERYKSLTP 239
 DB 231 IRDEVIAQVILQYKMLBELPTFTSYNSYFKDLKLANIKSGDETWQAKRVMERIVYAK 290
 QY 240 YPLSGENFR 248
 DB 291 IHLSEGNRR 299

RESULT 11

08CWP7 PRELIMINARY; PRT; 298 AA.

AC 08CWP7; PRELIMINARY; PRT; 298 AA.
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Primosome component (Helicase loader).
 GN DNAS OR SPR155.
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN NCBI_Taxid=171101;

RP SEQUENCE FROM N.A.

RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Albom W.B. Jr., Arnold J., Blazczak L.C., Burgett S., Dehoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C., LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsumura P., McMahon S.M., McHenry M., Mcleaster K., Mundy C.W., Niclas T.I., Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P., Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G., Zook C.A., Baltz R.H., Jaskunas S.R., Kosteck P.R. Jr., Skatrud P.L., Glaeser J.I.;
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6."
 RL J. Bacteriol. 183:5709-5717 (2001).
 DR EMBL; AE008523; AAL00359.1; -
 KW Helicase: Complete proteome.
 SQ SEQUENCE 298 AA; 34574 MW; 3EA8DBBD1DEDE6F22 CRC64;

Query Match 28.7%; Score 381; DB 16; Length 298;
 Best Local Similarity 34.7%; Pred. No. 7e-25;
 Matches 82; Conservative 50; Mismatches 100; Indels 4; Gaps 3;

QY 15 DCPNFKVGHVPELYVDNRIKIRYLQCPCKIYDEBRFAELITSHMQORDTLANKDI 74
 DB 65 DDDYIAKGYKPYLVNMHGVADVYSETPDLIAEKRAIKRLINIPSSSLKNSFLDV 124
 QY 75 VNMHRLDVMAADICTAINGEQVGLVYGPFGTKSFTLGAIANQLSKK-VRST 133
 DB 125 YKDDVQRLVLRMTIEPVNDYRN--NLKGLVYIGDPGVKSPVVALADLSERKGVST 182
 QY 134 IYLPDEFITLKGFGKSGFPEKRLHVRANILMLDDIGAEVTPWVRDEVIGPLIHYM 193
 DB 183 LHYSPFVIDVKNALSDGNVKTLDVEIKSEVLILLIDIGAEQSTWVRDEIIQVILQYMA 242
 QY 194 VHELPTFTSYNSADLERKWKATTKGSDETWQAKRVMERVAYLAKREPHLECANRR 248
 DB 243 QENLPTFTSYNSADLERKWKATTKGSDETWQAKRVMERIVYAKREPHLECANRR 298

RESULT 12

08CWP7 PRELIMINARY; PRT; 300 AA.

AC 08CWP7; PRELIMINARY; PRT; 300 AA.
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DE Hypothetical protein.
 GN DNAS OR GBS1668.
 OS Streptococcus agalactiae (serotype III).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 RN NCBI_Taxid=216495;
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEM316 / Serotype III;
 RX MEDLINE=2242508; PubMed=12354221;
 RA Glaeser P., Ruenick C., Buchrieser C., Chevalier F., Frangoul L., Wsadek T., Zouine M., Couve E., Lalloué L., Poyart C., Tieu-Cuot P., Kunet F.;
 RT "Genome sequence of Streptococcus agalactiae, a pathogen causing invasive neonatal disease."
 RL Mol. Microbiol. 45:1499-1513 (2002).
 DR EMBL; AL766852; CAD47327.1; -
 KW Sagalisc; gbs1668; -
 SQ SEQUENCE 300 AA; 34679 MW; 9COB4898661661B9 CRC64;

Query Match 27.5%; Score 365; DB 16; Length 300;
 Best Local Similarity 32.4%; Pred. No. 1.7e-23;
 Matches 77; Conservative 56; Mismatches 103; Indels 2; Gaps 2;

QY 12 KPADCPNFKVGHVPELYVDNRIKIRYLQCPCKIYDEBRFAELITSHMQORDTLANK 71
 DB 64 KKKDSQYIAKGYEPLVNMHGVADVSYLETRELIQAQKQAISDININVLKRSYRNIRM 123

QY 72 KDIYVNRDRDVAAMADICTAITNGEQVGLYVGPFGTGSFIIGAIANOLKSKK-V 130
 DB 124 TDFPINNESKMKMSQLDPEVETPSTVNH-KGLVLYGDMGVGKSYLMAANARELSERKGV 182
 QY 131 RSTIIVPEFIRTLKGGFKDGSFEKKLHVRREANIIMLDIGAEEVTPWVRDEVIGPLH 190
 DB 183 STTLHPSPFAIDVKNASISGTVKDEIDAVKSPILILDDIGAGQATSWVRDEILQVILQ 242
 QY 191 YRMVHELPTFPSSNFDVSELEHHLAMTRDGEETKKAIRIERYKSLSTPYLSEGNR 248
 DB 243 HRMLEELPTFTSNYSFNDLERKMANIKGSDETQAKRVMERVRVYLAIEPHLEGNRR 300

RESULT 13

Q8DY72 PRELIMINARY; PRT; 300 AA.
 AC Q8DY72;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Primosomal protein DnaI.
 GN DnaI OR SAG1621.
 OS Streptococcus agalactiae (serotype V).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=216466;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2603 V/R / Serotype V;
 RX MEDLINE=22222988; PubMed=12200547;
 RA Tetteijn H., Maatignat V., Cieslewicz M.J., Eissen J.A., Peterson S.,
 RA Messelt M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
 RA Medoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
 RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radune D., Fedotova N.B., Scanlan D., Khoult H., Mulligan S.,
 RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Jacobini E.T., Brettoni C., Galli G., Mariani M., Vegni P., Malone D.,
 RA Rinaldo D., Rappuoli R., Telford J.L., Kaper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V Streptococcus agalactiae";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL; AB014265; AAN00485.1; -.
 DR TIGR; SAG1621; -.
 KW Complete proteome.
 SQ SEQUENCE 300 AA; 34663 MW; 9BED36F86B36B395 CRC64;

Query Match 27.5%; Score 365; DB 16; Length 300;
 Best Local Similarity 32.4%; Pred. No. 1, 7e-23;
 Matches 77; Conservative 56; Mismatches 103; Indels 2; Gaps 2;

QY 12 KPADCPVGVGHVPELVVNNRIKIRYLOCPCKIKYDERFEALITSHMQRDTLNAKL 71
 DB 64 KNDSDOVIAGYEPFIVNMGYADVSYLRETRREIEAKQKAIIDRLNVLPRSYRIRM 123
 QY 72 KDIYVNRDRDVAAMADICTAITNGEQVGLYVGPFGTGSFIIGAIANOLKSKK-V 130
 DB 124 TDFPINNESKMKMSQLDPEVETPSTVNH-KGLVLYGDMGVGKSYLMAANARELSERKGV 182
 QY 131 RSTIIVPEFIRTLKGGFKDGSFEKKLHVRREANIIMLDIGAEEVTPWVRDEVIGPLH 190
 DB 183 STTLHPSPFAIDVKNASISGTVKDEIDAVKSPILILDDIGAGQATSWVRDEILQVILQ 242
 QY 191 YRMVHELPTFPSSNFDVSELEHHLAMTRDGEETKKAIRIERYKSLSTPYLSEGNR 248
 DB 243 HRMLEELPTFTSNYSFNDLERKMANIKGSDETQAKRVMERVRVYLAIEPHLEGNRR 300

RESULT 14
 Q97PC7 PRELIMINARY; PRT; 298 AA.
 ID Q97PC7;
 AC Q97PC7;

DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Primosomal protein DnaI.
 GN Spt711.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxID=1313;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TTG84;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tetteijn H., Nelson K.E., Paulsen I.T., Haft D.H., Dodson R.J.,
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtzapfel B., Khoult H., Wolf A.M., Utecherback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickenson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae";
 RL Science 293:498-506(2001).
 DR EMBL; AB007464; AAK75789.1; -.
 DR TIGR; SPI711; -.
 KW Complete proteome.
 SQ SEQUENCE 298 AA; 34259 MW; C4B5ECAF9C30059 CRC64;

Query Match 27.2%; Score 361; DB 16; Length 298;
 Best Local Similarity 34.7%; Pred. No. 3, 7e-23;
 Matches 83; Conservative 50; Mismatches 96; Indels 10; Gaps 4;

QY 15 DCNPFKGVGHVPELVVNNRIKIRYLOCPCKIKYDERFEALITSHMQRDTLNAKL 74
 DB 65 DTDYVAKGKYPFIVNMGYADVSYLRETRREIEAKQKAIIDRLNVLPRSYRIRM 124
 QY 75 YNMHRDRDVAAMADICTAITNGEQVGLYVGPFGTGSFIIGAIANOLKSKK-V 130
 DB 125 TDFPINNESKMKMSQLDPEVETPSTVNH-KGLVLYGDMGVGKSYLMAANARELSERKGV 179
 QY 131 RSTIIVPEFIRTLKGGFKDGSFEKKLHVRREANIIMLDIGAEEVTPWVRDEVIGPLH 190
 DB 180 STTLHPSPFAIDVKNASISGTVKDEIDAVKSPILILDDIGAGQATSWVRDEILQVILQ 239
 QY 191 YRMVHELPTFPSSNFDVSELEHHLAMTRDGEETKKAIRIERYKSLSTPYLSEGNR 248
 DB 240 YRMQENLPTFTSNYSFNDLERKMANIKGSDETQAKRVMERVRVYLAIEPHLEGNRR 298

RESULT 15

O9CH9 PRELIMINARY; PRT; 293 AA.
 ID O9CH9;
 AC O9CH9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Primosomal protein DnaI.
 GN DnaI OR L10752.
 OS Streptococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 NCBI_TaxID=1360;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TL1403;
 RX MEDLINE=2135186; PubMed=11337471;
 RA Bolotin A., Wincker P., Mauer S., Jallion O., Malarre K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis TL1403";
 RL Genome Res. 11:731-753(2001).
 DR EMBL; AE006308; AAK04850.1; -.
 DR InterPro; IPR003593; AAA_Atpase.

DR SMART; SM00382; AAA; 1.
KM Complete proteome.
SQ SEQUENCE 293 AA; 33904 MW; D75C530A69445B3 CRC64;

Query Match 25.8%; Score 342; DB 16; Length 293;
Best Local Similarity 33.3%; Pred. No. 1.6e-21;
Matches 85; Conservative 49; Mismatches 97; Indels 24; Gaps 6;

```
QY 2 KQOQKHVDGKHFADCFNFKGHVPELYVNNRIKIRYLQCPCKIKYDERFPAELITSHH 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 57 REHEKEKEKEKRA-----ADGYEPVLINHGADVS---QTNELAQQAQNLRRRN 108
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 ---MORDTLNAKLDIYMAHRDLDAVMAADICTAITNGEQVKGLYLGPFGTGSFIL 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 109 IIGLPKDLQVTLADLDVQRIKRYQALVDF---ITNPKKGGLYLGDFGVGKSPVL 165
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 GAIANQLKSKKQVSTIIYLPPIRTILKGFQDGSPEKK--LHRVREANILMLDDIGAEE 175
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 166 AAMANELAKKGISTILLHYPTFI-----SDLDFDNAAKVMVEIKASQVLVLDIGAEQ 218
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 VTPWVDEVIIGPLIHRMHELPFFSSNFDYSELEHILAMTRDGEKTKAARIIERVKS 235
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 219 NNAWVDSILQVLIQHRMQENLPTPTSNI, RMELEQLAETGRADEIWPAKRVMERVKY 278
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 236 LSTPYFLSGENPRNN 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 LAEMRLBGTNRHD 293
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: December 15, 2003, 15:15:58
Job time : 29.8858 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 45.6369 Seconds
(without alignments)
1088.625 Million cell updates/sec

Title: US-09-689-952-2
Perfect score: 1643
Sequence: 1 MCGGSGIMKQFISINTSQD.....ERVKSLSTPYPLSGENFRNN 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq.19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1643	100.0	313	22	AA847317
2	1378	83.9	307	22	ABP38947
3	1353	82.3	299	22	AA682049
4	1353	82.3	299	22	AA682337
5	1328	80.8	250	22	AA847316
6	851	51.8	164	22	AA847315
7	617.5	37.6	155	22	AA681811
8	552.5	33.6	307	23	AB849355
9	429.5	26.1	300	23	ABP27829

10	403.5	24.6	293	23	AB854065
11	393	23.9	92	23	ABP09981
12	393	23.9	300	23	ABP27828
13	392.5	23.9	298	24	ABU02187
14	112	6.8	291	22	AB853757
15	110	6.7	866	22	AAU38917
16	110	6.7	872	20	AAV34745
17	109.5	6.7	294	23	AB854759
18	109	6.6	630	23	AB848390
19	108.5	6.6	400	21	AA634347
20	107	6.5	331	23	ABP65341
21	107	6.5	331	23	ABP65354
22	107	6.5	331	23	ABP65547
23	106.5	6.5	797	21	AA181819
24	106.5	6.5	897	22	AB862660
25	106	6.5	435	21	AA634039
26	106	6.5	500	21	AA634038
27	105	6.4	248	22	AA698965
28	105	6.4	479	22	ABG18303
29	105	6.4	798	22	AA696715
30	105	6.4	1886	19	AAW54241
31	104.5	6.4	386	22	ABG25088
32	104	6.3	399	21	AA626569
33	103.5	6.3	397	22	AB859915
34	103.5	6.3	399	21	AA646306
35	103	6.3	308	24	ABU01560
36	102.5	6.2	304	23	ABP40308
37	102	6.2	389	24	ABJ25711
38	102	6.2	389	24	ABJ26311
39	102	6.2	453	23	ABP30561
40	102	6.2	484	23	ABP27967
41	102	6.2	1935	23	ABG31649
42	101.5	6.2	784	22	ABG30180
43	101.5	6.2	827	22	ABP73784
44	101	6.1	451	24	ABJ26363
45	101	6.1	2539	21	AA181819

ALIGNMENTS

RESULT 1	AA847317	standard; Protein; 313 AA.
ID	AA847317;	
AC	AA847317;	
XX		
DT	29-AUG-2001	(first entry)
XX		
DE	S. aureus DnaI.	
XX		
KW	DnaI; S. aureus; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;	
KW	screening assay.	
XX		
OS	Staphylococcus aureus.	
XX		
PN	WO200146383-A2.	
XX		
PD	28-JUN-2001.	
XX		
FP	21-DEC-2000; 2000WO-US35180.	
XX		
PR	22-DEC-1999; 99US-0470512.	
XX		
PR	12-OCT-2000; 2000US-0689952.	
XX		
PA	(PHAG-) PHAGETECH INC.	
XX	(WILLI) WILLIAMS K M.	
XX		
PI	Pelletier J, Gros P, Dubow M;	
XX		
XX	WPI; 2001-418052/44.	
DR	N-PSDB; AAC86105.	
XX		

Lactococcus lactis
Human ORF protein
Streptococcus poly
S. pneumoniae type
Lactococcus lactis
C. pneumoniae CTII
C. pneumoniae prot
Lactococcus lactis
Listeria monocytog
Zea mays protein f
Bifidobacterium lo
Bifidobacterium lo
Plasmodium falcipar
Drosophila melanog
Zea mays protein f
Zea mays growth and
E. coli growth and
Novel human diagno
Putative P. abyssi
Rattus norvegicus
Novel human diagno
Arabidopsis thaliana
Drosophila melanog
Arabidopsis thaliana
S. pneumoniae type
Staphylococcus epi
Aspergillus fumiga
Aspergillus fumiga
Streptococcus poly
Streptococcus poly
Amino acid distrib
Novel human diagno
Candida albicans e
Aspergillus fumiga
Plasmodium falcipar

PT Novel DnaI polypeptides useful for treating and diagnosing microbial,
PT preferably bacterial, diseases such as those caused by *Staphylococcus*
XX aureus -
XX
XX Disclosure, Fig 1, 107bp, English.
XX
XX This sequence shows DnaI derived from *S. aureus*. *S. aureus* DnaI
CC interacts with a growth-inhibitory bacteriophage 77 ORF 104
CC gene product, to form the basis of a screening assay. DnaI
CC polypeptides and polynucleotides are useful for treating microbial,
CC preferably bacterial, especially *Staphylococcus*, infections. DnaI
CC polypeptides and polynucleotides are useful for biological, diagnostic,
CC prophylactic, clinical and therapeutic use, and as components in
CC databases useful for search analyses as well as in sequence analysis
CC algorithms.
XX
XX Sequence 313 AA,
SQ

Query Match 100.0%; Score 1643; DB 22; Length 313;
Best Local Similarity 100.0%; Pred. No. 2,5e-139;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSGQSIMKQFKSIINTSDPEKRIEIKKEVINDPVKQFLERHRELTMAMIDEDLVN 60
DB 1 MGSGQSIMKQFKSIINTSDPEKRIEIKKEVINDPVKQFLERHRELTMAMIDEDLVN 60
QY 61 LQRYKDDQKHYDGHKPADCPNFKVGHVPELYVNNRIRIKRYLOCPCKIXYDEERPEAEI 120
DB 61 LQRYKDDQKHYDGHKPADCPNFKVGHVPELYVNNRIRIKRYLOCPCKIXYDEERPEAEI 120
QY 121 TSHHMQRDTLANAKLKOIYNNHRDLVMAADICTAITNGEQVKGILYGFPGTGSFPI 180
DB 121 TSHHMQRDTLANAKLKOIYNNHRDLVMAADICTAITNGEQVKGILYGFPGTGSFPI 180
QY 181 LGAIANOLSKKQVRSITIIYLPFIRITLKGFGDGSFEKKLHVRVREANITLMDIGAEV 240
DB 181 LGAIANOLSKKQVRSITIIYLPFIRITLKGFGDGSFEKKLHVRVREANITLMDIGAEV 240
QY 241 PWRDEVIGPLHYRMVHELPTFSSNPDYSELEHILAMTRDGEKTKARITIERVSL 300
DB 241 PWRDEVIGPLHYRMVHELPTFSSNPDYSELEHILAMTRDGEKTKARITIERVSL 300
QY 301 TPYFLSGENFRNN 313
DB 301 TPYFLSGENFRNN 313

RESULT 2
ABP38947
ID ABP38947 standard; Protein; 307 AA.

AC ABP38947;

DT 24-JUL-2002 (first entry)

DE *Staphylococcus epidermidis* ORF amino acid sequence SEQ ID NO:3792.

KW *Staphylococcus epidermidis*, open reading frame; ORF; bacterial infection;
anti-bacterial; gene therapy.

OS *Staphylococcus epidermidis*.

PN US6380370-B1.

PD 30-APR-2002.

PF 13-AUG-1998; 98US-0134001.

PR 14-AUG-1997; 97US-055779P.

XX 08-NOV-1997; 97US-064964P.

XX (GENO-) GENOME THERAPEUTICS CORP.

PI Doucette-Stamm LA, Bush D;
XX
XX WPI: 2002-381255/41.
DR N-PSDB; ABN91492.

PT Novel isolated nucleic acid encoding a *Staphylococcus epidermidis*
XX polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX Disclosure; SEQ ID 3792; 267bp; English.

CC ABN90538 to ABN93374 represent *Staphylococcus epidermidis* open reading
CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC given in ABP35124 to ABP37960. The *S. epidermidis* sequences have
CC antibacterial activity and can be used in gene therapy. The sequences
CC can also be used in the diagnosis and treatment of bacterial infections,
CC particularly *S. epidermidis* infections. The sequences can be used to
CC screen for compounds able to interfere with the *S. epidermidis* life
CC cycle or inhibit *S. epidermidis* infection.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site.

XX
XX Sequence 307 AA;
SQ

Query Match 83.9%; Score 1378; DB 23; Length 307;
Best Local Similarity 82.7%; Pred. No. 1,6e-115;
Matches 254; Conservative 34; Mismatches 19; Indels 0; Gaps 0;

QY 7 IMKQPSIINTSDPEKRIEIKKEVINDPVKQFLERHRELTMAMIDEDLVNLOEKYD 66
DB 1 LMKSPFINMDSQNLDR1QIKQVNIYNDVGHFLEKRSNITNEMIDEDLVNLOEKYD 60
QY 67 QQGVYDGHKPADCPNFKVGHVPELYVNNRIRIKRYLOCPCKIXYDEERPEAEI 126
DB 67 QQGVYDGHKPADCPNFKVGHVPELYVNNRIRIKRYLOCPCKIXYDEERPEAEI 120
QY 127 RDTLANAKLKOIYNNHRDLVMAADICTAITNGEQVKGILYGFPGTGSFPI 186
DB 127 RDTLANAKLKOIYNNHRDLVMAADICTAITNGEQVKGILYGFPGTGSFPI 180
QY 187 QLSKQVRSITIIYLPFIRITLKGFGDGSFEKKLHVRVREANITLMDIGAEV 246
DB 187 QLSKQVRSITIIYLPFIRITLKGFGDGSFEKKLHVRVREANITLMDIGAEV 240
QY 247 VIGPLHYRMVHELPTFSSNPDYSELEHILAMTRDGEKTKARITIERVSL 306
DB 247 VIGPLHYRMVHELPTFSSNPDYSELEHILAMTRDGEKTKARITIERVSL 300
QY 307 GENFRNN 313
DB 307 GENFRNN 313

RESULT 3
AAG82049
ID AAG82049 standard; Protein; 299 AA.

AC AAG82049;

DT 03-SEP-2001 (first entry)

DE *S. epidermidis* open reading frame protein sequence SEQ ID NO:1192.

KW *Staphylococcus epidermidis* SRI strain; infection; diagnosis;
vaccination; endocarditis.

OS *Staphylococcus epidermidis*.

PN WO200134809-A2.

PD 17-MAY-2001.

XX 09-NOV-2000; 2000WO-US30782.


```
RESULT 5
AAB47316
ID AAB47316 standard; Protein; 250 AA.
XX
AC AAB47316;
XX
DT 29-AUG-2001 (first entry)
XX
DE Amino acids 64-313 of S. aureus DnaI.
XX
KM DnaI; S. aureus; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
KW screening assay.
XX
OS Staphylococcus aureus.
XX
PN MO200146383-A2.
XX
PD 28-JUN-2001.
XX
PF 21-DEC-2000; 2000MO-US35180.
XX
PR 22-DEC-1999; 99US-0470512.
XX
PR 12-OCT-2000; 2000US-0689952.
XX
PA (PHAG-) PHAGESTEC INC.
XX
PA (WILL/) WILLIAMS K M.
XX
PI Pelletier J, Gros P, Dubow M;
XX
DR WPI; 2001-418052/44.
XX
PT Novel DnaI polypeptides useful for treating and diagnosing microbial,
PT preferably bacterial, diseases such as those caused by Staphylococcus
PT aureus -
XX
PS Disclosure; Fig 15; 107pp; English.
XX
CC This sequence shows a DnaI polypeptide derived from S. aureus. S.
CC aureus DnaI interacts with a growth-inhibitory bacteriophage 77 ORF
CC 104 gene product, to form the basis of a screening assay. DnaI
CC polypeptides and polynucleotides are useful for treating microbial,
CC preferably bacterial, especially Staphylococcus, infections. DnaI
CC polypeptides and polynucleotides are useful for biological, diagnostic,
CC prophylactic, clinical and therapeutic use, and as components in
CC databases useful for search analyses as well as in sequence analysis
CC algorithms.
XX
SQ Sequence 250 AA;
XX
Query Match 80.8%; Score 1328; DB 22; Length 250;
Best Local Similarity 100.0%; Pred. No. 3,7e-111;
Matches 250; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 64 YKQOQKHGDKFPADCFNFKGVHVELYDNNRIKIRYLOCPCKIKYDEERFEAEITSH 123
DB 1 YKQOQKHGDKFPADCFNFKGVHVELYDNNRIKIRYLOCPCKIKYDEERFEAEITSH 60
QY 124 HMQRDITLNAKLKDIYNNHRRRLDYVMAADICTAITNGEOVKGLYLYGPGTGKSFILGA 183
DB 61 HMQRDITLNAKLKDIYNNHRRRLDYVMAADICTAITNGEOVKGLYLYGPGTGKSFILGA 120
QY 184 IANOLSKKVRSTIIYLPFIRTLKGGFKDGSFEKTLHVRREANITLMDDIGAEVTPWV 243
DB 121 IANOLSKKVRSTIIYLPFIRTLKGGFKDGSFEKTLHVRREANITLMDDIGAEVTPWV 180
QY 244 RDEVIGPLHYRWVHELPTFFSSNFDYSELHHLAMTRDGEKTKAARIIRVKSISTPY 303
DB 181 RDEVIGPLHYRWVHELPTFFSSNFDYSELHHLAMTRDGEKTKAARIIRVKSISTPY 240
QY 304 FLGGENFRNN 313
DB 241 FLGGENFRNN 250
```

```
RESULT 6
AAB47315
ID AAB47315 standard; Protein; 164 AA.
XX
AC AAB47315;
XX
DT 29-AUG-2001 (first entry)
XX
DE Amino acids 150-313 of S. aureus DnaI.
XX
KM DnaI; S. aureus; inhibitor; bacteriophage 77; ORF 104; phage 77ORF104;
KW screening assay.
XX
OS Staphylococcus aureus.
XX
PN MO200146383-A2.
XX
PD 28-JUN-2001.
XX
PF 21-DEC-2000; 2000MO-US35180.
XX
PR 22-DEC-1999; 99US-0470512.
XX
PR 12-OCT-2000; 2000US-0689952.
XX
PA (PHAG-) PHAGESTEC INC.
XX
PA (WILL/) WILLIAMS K M.
XX
PI Pelletier J, Gros P, Dubow M;
XX
DR WPI; 2001-418052/44.
XX
DR N-PSDB; AAC6104.
XX
PT Novel DnaI polypeptides useful for treating and diagnosing microbial,
PT preferably bacterial, diseases such as those caused by Staphylococcus
PT aureus -
XX
PS Claim 41; Fig 15; 107pp; English.
XX
CC This sequence shows a DnaI polypeptide derived from S. aureus. S.
CC aureus DnaI interacts with a growth-inhibitory bacteriophage 77 ORF
CC 104 gene product, to form the basis of a screening assay. DnaI
CC polypeptides and polynucleotides are useful for treating microbial,
CC preferably bacterial, especially Staphylococcus, infections. DnaI
CC polypeptides and polynucleotides are useful for biological, diagnostic,
CC prophylactic, clinical and therapeutic use, and as components in
CC databases useful for search analyses as well as in sequence analysis
CC algorithms.
XX
SQ Sequence 164 AA;
XX
Query Match 51.8%; Score 851; DB 22; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.6e-68;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 150 AADDICTAITNGEOVKGLYLYGPGTGKSFILGAIANOLSKKVRSTIIYLPFIRTLK 209
DB 1 AADDICTAITNGEOVKGLYLYGPGTGKSFILGAIANOLSKKVRSTIIYLPFIRTLK 60
QY 210 GFKDGSFEKTLHVRREANITLMDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFD 269
DB 61 GFKDGSFEKTLHVRREANITLMDDIGAEVTPWVRDEVIGPLHYRWVHELPTFFSSNFD 120
QY 270 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFLGGENFRNN 313
DB 121 YSELHHLAMTRDGEKTKAARIIRVKSISTPYFLGGENFRNN 164
XX
RESULT 7
AAG81811
ID AAG81811 standard; Protein; 155 AA.
XX
AC AAG81811;
```

XX 03-SEP-2001 (first entry)
 DT S. epidermidis open reading frame protein sequence SEQ ID NO:716.
 DE Staphylococcus epidermidis SRI strain; infection; diagnosis;
 KW vaccination; endocarditis.
 XX Staphylococcus epidermidis.
 OS MO200134809-A2.
 PN 17-MAY-2001.
 PD 09-NOV-2000; 2000MO-US30782.
 PF 09-NOV-1999; 99US-0164258.
 PR (GLAXO) GLAXO GROUP LTD.
 PA
 XX Kimmberly WJ;
 PI WPI, 2001-316495/33.
 DR N-PSDB; AAH52661.
 XX
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 useful for vaccinating against infections, e.g. endocarditis -
 Claim 18; Page 223; 2188bp; English.
 XX
 CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 155 AA;
 Query Match 37.6%; Score 617.5; DB 22; Length 155;
 Best Local Similarity 77.0%; Pred. No. 1,4e-47;
 Matches 114; Conservative 19; Mismatches 12; Indels 3; Gaps 1;
 QY 18 SQDFEKRIEIKKEVINDDPVKQFLFAHRAELTNAMIDEDLVLYQEKYQOQKHYYDGHKA 77
 DB 4 SQMLDRRIQIKQNVINDTVKAFLEKNRSENITNEMIDEDLVLYQEKYQOQKHYYDGHKA 63
 QY 78 DCNPFKGVHPELVYDNNRIRKIRYLOCPCKIKYDERFAELITSHHMQDITLNAKIKOI 137
 DB 64 DCNPFKGVHPELVYDNNRIRKIRYLOCPCKIKYDERFAELITSHHMQDITLNAKIKOI 123
 QY 138 YNNHRDLVAVMAADICTAITNGEQVK 165
 DB 124 YNNHRDLVAVMAADICTAITNGEQVK 148
 RESULT 8
 ABB49355
 ID ABB49355 standard; Protein; 307 AA.

XX ABB49355;
 AC 05-FEB-2002 (first entry)
 DT Listeria monocytogenes protein #2059.
 DE Listeria monocytogenes protein #2059.
 KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KM vitamin B12; bacterial infection; disease.
 XX Listeria monocytogenes.
 OS MO200177335-A2.
 PN 16-OCT-2001.
 PD 11-APR-2001; 2001MO-FR01118.
 PF 11-APR-2000; 2000FR-0004629.
 PR (INSP) INST PASTEUR.
 PA
 XX Buchrieser C, Frangeul L, Couve E, Ruenliok C, Faini H, Dehoux P,
 PI Dussurgeat O, Chetoui F, Nedjati H, Glaeser P, Kunet F, Cossart P,
 PI Daniels J, Goebel W, Krefe J, Kuhn M, Ng E, Vazquez-Boland JA,
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A,
 PI Chakraborty T, Domann E, Hain T, Berche P, Chardit A, Durant L,
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N,
 PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J,
 PI Rose M, Voss H;
 DR WPI, 2002-010914/01.
 XX
 PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides -
 XX
 PS Claim 6; SEQ ID NO 2060; 192bp; French.
 XX
 CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccines compositions for the treatment or prevention of infections by L.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 307 AA;
 Query Match 33.6%; Score 552.5; DB 23; Length 307;
 Best Local Similarity 37.8%; Pred. No. 2,4e-41;
 Matches 113; Conservative 69; Mismatches 104; Indels 13; Gaps 4;
 QY 19 ODPEKRIEIKKEVINDDPVKQFLFAHRAELTNAMIDEDLVLYQEKYQOQKHYYDGHKPAD 78
 DB 15 RDPFEKRYQGLKQVLYHYPIDQFFKHKKEVLEQVNLSTLYEPMTHKKFTGEQ--- 71
 QY 79 CPNPFKGVHPELVYDNNRIRKIRYLOCPCKIKYDERFAELITSHHMQDITLNAKIKOI 138
 DB 72 -ETLMPGVAPKLVINGEFTVYTPYFKKEIKEDKRAVRRIRSLYMPKQVVDANLADPY 130
 QY 139 YNNHRDLVAVMAADICTAITN-----GQVAGVLYGFGTGSKPIIGAIANQLSKKV 193

Db 131 TDESR---QLATVEAVQFLNNYPPKSGERYKGLFIHGSFETGKSYLLGALAKELAKGI 187
 QY 194 RSTIIYPERFIRTLKGGFKDGSPEKRLHRYEANIIMLDIGAEVTPWPRDEVITGPI 253
 Db 188 STIIYVPERFIRTLKGGFKDGSPEKRLHRYEANIIMLDIGAEVTPWPRDEVITGPI 247
 QY 254 YRMVHELPTFFSNFDYSELEHHLAMTRDG-EETKARILIERVKSISTPYFLSGENP 311
 Db 248 FRMGELPTFFSNFDYSELEHHLAMTRDG-EETKARILIERVKSISTPYFLSGENP 306

RESULT 9

ABP27829 ID ABP27829 standard; Protein; 300 AA.

AC ABB27829;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4834.

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;

KM group A streptococcus; Streptococcus pyogenes; antibacterial;

XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus pyogenes.

PN W0200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB04789.

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Telford H;

DR WPI; 2002-352536/38.

DR N-PSDB; ABB68460.

XX New Streptococcus protein for the treatment or prevention of infection

PT or disease caused by Streptococcus bacteria, such as meningitis, and

PT for detecting a compound that binds to the protein -

XX Claim 1; Page 3645; 4525pp; English.

XX The invention relates to a protein (ABP25413-ABP30895) from group B

CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS

CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in

CC the specification. The proteins have antibacterial and antiinflammatory

CC activity. (I), nucleic acids encoding (I), ABB6044-ABN71526 and

CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by

CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a

CC biological sample. (I) is used to determine whether a compound binds to

CC used as a vaccine or diagnostic composition. The disease caused by

CC Streptococcus that is prevented or treated may be meningitis. Nucleic

CC acid encoding (I) may be used to recombinantly produce (I) and may be

CC used in gene therapy. Antibodies to (I) are used for affinity

CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.

XX Sequence 300 AA;

Query Match 26.1%; Score 429.5; DB 23; Length 300;
 Best Local Similarity 31.2%; Pred. No. 2,7e-30;
 Matches 94; Conservative 74; Mismatches 126; Indels 7; Gaps 4;
 QY 12 KSIIINSQDEPERIKKEKKEVINDDPKQFLHARAEITAMIDEDLVNQEYKDOQKH 71
 Db 6 ETWAKLGQONTRVNSDQIQTILADPEVASFISQH-HLSGQINTLSKRNQFLVERQKY 63
 QY 72 DGKHPADCFNFAVHVELVYDNNRIKIRYLOCPCKIKYEEEPFELTSHMQRDITLN 131
 Db 64 ---QKDPSTIAGKYDPIAMNEGADVSYLEKELVEAKQAISRIQVLSPKSYRH 120
 QY 132 AKLDIYMHNRDLVDVMAADDICTAITNGEQVKGILYGPFGTGSFTIGAIAMQSK 191
 Db 121 IHLSDIDVNNASWAEAFSALIDFVEQPSAEQ-KGLYLYDMDGIGKSYLLAAAHLESEK 179
 QY 192 K-VRSITIIYPERFIRTLKGGFKDGSPEKRLHRYEANIIMLDIGAEVTPWPRDEVITG 250
 Db 180 KGVSTIIHFPSPFAIDKNAISNGSVKBEIDAVKVPVLLIDDIGAEQATSWRDEVITQV 239
 QY 251 LILHYRMVHELPTFFSNFDYSELEHHLAMTRDGEETKARILIERVKSISTPYFLSGENP 310
 Db 240 ILQYRMHELPTFFSNFDYSELEHHLAMTRDGEETKARILIERVKSISTPYFLSGENP 299
 QY 311 R 311
 Db 300 R 300

RESULT 10

ABBS4065 ID ABBS4065 standard; Protein; 293 AA.

AC ABB54065;

DT 16-MAY-2002 (first entry)

DE Lactococcus lactis protein dnaI.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

OS Lactococcus lactis IL1403.

PN FR2807446-A1.

PD 12-OCT-2001.

PF 11-APR-2000; 2000FR-0004630.

PR 11-APR-2000; 2000FR-0004630.

PA (INRG) INRA INST NAT RECH AGRONOMIQUE.

PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;

PI WPI; 2002-043418/06.

PT New nucleotide sequence useful in the identification or Lactococcus

PT lactis and related species -

XX Claim 6; SEQ ID No 767; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide

CC sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The

CC nucleic acid sequence is useful in the detection and/or amplification of

CC related species. The proteins of the invention are useful for the

CC biosynthesis or biodegradation of a composition of interest. The

CC invention helps research in lactic bacteria, particularly useful in the

CC production of yogurt and cheese.

CC Note: The sequence data for this patent is based on equivalent patent

CC W020017734 (published 18-Oct-2001) which is available in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX	Sequence	293 AA:
XX	Query Match	24.6%; Score 403.5; DB 23; Length 293;
XX	Best Local Similarity	32.3%; Pred. No. 5,6e-28;
XX	Matches 101; Conservative	61; Mismatches 124; Indels 27; Gaps 8;
QY	8 MKQFSLTINTSODPEFKRIEIKKEVYINDPDYQFLEAHRALETLNMIMEDLNVLOEY-KD 66	
DB	1 MESIGDILGKRQDIRENFEKVLVAEVLKNADVOAFIAQH--QMTSDEIGRSYSKEFYEVNE 58	
QY	67 QOKHYDGKFAFDCPFVFGKGVHVELYVDNRIRIIRYLQCECKIKYDEERFEALITSSH-- 124	
DB	59 HEKEFEKGRKA-----ADGYEPEVLIMNHGYADVST---QTMELAQQAQVLLRRMNI 110	
QY	125 -MQRDTLNAKLKDIYMNHRDLVDVMAADDICTAITNGEQVGLYLYGFGTGKSPITIGA 183	
DB	111 GLPKOLKQVTLADIALDDVQRIKPYQALVDF---ITNPFKKKGLYLGDFGVGKSFMLAA 167	
QY	184 IANQLSKSKVASTIYLYPEPFRITLKGFGDGSFEKK--LHVRRANILMLDIDGAEVY 240	
DB	168 MANELAKKGISTTLTLPYLPFI-----SDLPFNDAKVMWNKIKASQVYLVDIDIAEQNN 220	
QY	241 FAVRDEVIGPLHYRMVHMLPTFPSSNPDYSELEHNLAMTRDGEKTKAARIIEVKSL 300	
DB	221 AWRDSTIQVILQHMQENLPTFPISNLMELBOHLATKRADEIWPAKRMERVKYLA 280	
QY	301 TPYPLSGENFRNN 313	
DB	281 EEMRLEGTRRHMD 293	
XX	RESULT 11	
XX	ABP09981	
XX	ID ABP0981 standard; Protein; 92 AA.	
XX	ABP09981;	
XX	25-JUN-2002 (first entry)	
DE	Human ORFX protein sequence SEQ ID NO:19944.	
XX	Human, open reading frame; ORFX; gene therapy; cancer; cirrhosis;	
XX	hyperproliferative disorder; psoriasis; benign tumor; haemorrhage;	
XX	degenerative disorder; osteoarthritis; neurodegenerative disorder;	
XX	cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;	
XX	hyperension; hypothyroidism; cholesterol ester storage disease;	
XX	immune deficiency; immune disorder; infectious disease;	
XX	autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;	
XX	myasthenia gravis.	
XX	Homo sapiens.	
XX	OS	
XX	WO200192523-A2.	
XX	PD	
XX	06-DEC-2001.	
XX	29-MAY-2001; 2001WO-US10836.	
XX	30-MAY-2000; 2000US-206132P.	
XX	29-AUG-2000; 2000US-228716P.	
XX	(CURA-) CURAGEN CORP.	
XX	Shimkets RA, Leach MD;	
XX	WPI; 2002-106308/14.	
XX	N-PSDB; ABN25733.	
XX	Novel human polypeptides and polynucleotides useful for diagnosing,	
XX	preventing and treating cardiovascular disease, neurodegenerative,	
XX	hyperproliferative disorders and autoimmune disorders	

PS Disclosure; SEQ ID 19944; 1037bp; English.

XX The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORF, where X is 1-11491 (see Table 1
CC in the specification). ABRN15762 to ABRN27252 encode the human ORF
CC proteins given in ABRN00010 to ABRN1500. ORF proteins are useful for
CC treating or preventing a pathology associated with an ORF-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORF-associated disorder. ORF polynucleotide
CC sequences can be used in gene therapy. ORF sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumors, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORF proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.

CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
XX
SQ Sequence 92 AA;

Query Match 23.9%; Score 393; DB 23; Length 92;
Beet Local Similarity 81.3%; Pred. No. 1.1e-27;
Matches 74; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

OY 163 QVKKLYYGPFGTSGKSPFLGAINANOLSKKVRSTIYLPFIRPLTKGSPKDGSEPKLHR 222
Db 2 RKGLYIHGFGTKSPILKSIANOLSKSKVPSITVLPYIRLTKGSPKDGSTNELER 61
OY 223 VREANILMLDDIGAEEVTPWVRDEVIGPLH 253
Db 62 IREANILMLDDIGAEDVTWVRDEVIGPLH 92

RESULT 12
ID ABR27828 standard; Protein; 300 AA.
XX ABR27828;
XX
XX ABR27828;
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 4832.
XX
KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
XX MO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
XX 24-NOV-2000; 2000GB-0028727.
XX 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX
XX Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Tectelin H;
 XX WPI; 2002-352536/38.
 DR N-PSDB; ABN6459.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 PS Claim 1; Page 3645; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and anti-inflammatory
 CC activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
 CC antibodies that bind (1) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (1) are used to detect Streptococcus in a
 CC biological sample. (1) is used to determine whether a compound binds to
 CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (1) may be used to recombinantly produce (1) and may be
 CC used in gene therapy. Antibodies to (1) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 SQ Sequence 300 AA;
 Query Match 23.9%; Score 393; DB 23; Length 300;
 Best Local Similarity 28.5%; Pred. No. 5,1e-27;
 Matches 88; Conservative 80; Mismatches 127; Indels 14; Gaps 5;
 QY 4 GGSIMKQFKSIINTSODPEKRIEIKKEVINDPVKQPLAHRALTNAMIDEDLVLOE 63
 DB 5 GQALNENGRVBRNTN-----DELQIMLADQVAFIKTH--QLSOREINISMSKFNQ 55
 QY 64 YKQKQKHYDGHKRPDCNPFKGNHVELYVNNRKIKIYLLQCPCKIKYDEREFRELITSH 123
 DB 56 FLIERKRF---KNKDSQYIAKGYEPILVMEGVADYLLTREILIEKQKQALSDRLNLV 112
 QY 124 HMORDTLNAKIKDIYNNHRDLVMAADICTAITNGEYQKGLYLGPFGTGKSFILGA 183
 DB 113 NLRPSYNNIMTDPDINNESRMKMSQLDFVEYEPYNN-KGLYLYGDMGVCKSYLMAA 171
 QY 184 IANQLSKK--VRSITVILPEPITLKGKFGDSFEKTLHVRNANIIMDDIGAEVTPW 242
 DB 172 MARELSERKGVSTLHFPFAIDVKNALISSTGYKDEIDAVKSVPIILDDIGAEQXTSW 231
 QY 243 VREVGIGPLAHYNNVHELPFSSNPYSLNHLNLTNRGSEETKAARIIEYKSLSTR 302
 DB 232 VREBIIQVLIQHRLLEELPTFTSNYSFNDLERKNANIKGSDTQAKRVERRYLAIE 291
 QY 303 YPLSGENFR 311
 DB 292 PHLEGPNRR 300
 RESULT 13
 ABU02187
 ID ABU02187 standard; Protein; 298 AA.
 XX
 AC ABU02187;
 DT 11-FEB-2003 (first entry)
 XX
 DB S. pneumoniae type 4 strain protein from coding region #1764.
 XX
 KW Bacterial meningitis; pneumonia; sepsis; otitis media;
 KW ear infection; antiinflammatory; antibacterial; immunostimulant;
 KW auditory; respiratory; gene therapy; vaccine.

XX
 OS Streptococcus pneumoniae type 4 strain.
 XX
 PN NO200277021-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 27-MAR-2002; 2002MO-IB02163.
 XX
 PR 27-MAR-2001; 2001GB-0007658.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PI Maignani V, Tectelin H, Fraser C;
 DR WPI; 2003-040579/03.
 DR N-PSDB; ABX07476.
 XX
 PT New proteins and nucleic acid molecules from Streptococcus pneumoniae,
 PT useful as medicaments for treating or preventing a disease or infection
 PT due to streptococcus bacteria, such as pneumonia, sepsis, otitis media
 PT or ear infection -
 PS Claim 1; SEQ ID NO 3528; 56pp; English.
 XX
 CC The invention relates to a protein comprising or having at least 50%
 CC identity to any of the 2469 amino acid sequences, identified in the
 CC specification (available on a computer readable format), or its fragment,
 CC expressed from 2469 of 2489 identified DNA coding regions from the
 CC Streptococcus pneumoniae type 4 strain genomic sequence appearing as
 CC ABS5454. Also included are an antibody which binds one of the
 CC proteins, treating a patient by administering the protein, DNA or
 CC antibody (in a composition), a kit comprising first and second primers,
 CC which are the nucleic acid cited above or fragments between nucleotides
 CC 8-100 of a sequence not defined in the specification, for amplifying a
 CC target sequence contained within a Streptococcus nucleic acid sequence,
 CC where the first primer is substantially complementary to the target
 CC sequence and the second primer is substantially complementary to the
 CC complement of the target sequence, and where the parts of the primers
 CC having substantial complementarity define the terminal of the target
 CC sequence to be amplified, assay comprising contacting a test compound
 CC with the protein, and determining whether the test compound binds to the
 CC protein and a Streptococcus pneumoniae bacterium, where one or more
 CC genes encoding the proteins has been rendered inactive. The proteins,
 CC nucleic acid molecules, antibody and compositions are useful as
 CC medicaments for treating or preventing a disease or infection due to
 CC streptococcus bacteria, particularly S. pneumoniae, such as pneumonia,
 CC sepsis, otitis media or ear infection. They are also useful in developing
 CC vaccines, diagnostics and antibiotics. The methods are useful for
 CC identifying immunodominant proteins. The present sequence is one of
 CC the 2469 proteins expressed by the identified coding regions from the
 CC genomic sequence.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 298 AA;
 Query Match 23.9%; Score 392.5; DB 24; Length 298;
 Best Local Similarity 32.3%; Pred. No. 5,6e-27;
 Matches 96; Conservative 64; Mismatches 116; Indels 21; Gaps 8;
 QY 21 FEKRIEIKKEVINDPVKQPLAHRALTNAMIDEDLVLOEY-KDQKHYDGHKRPDC 79
 DB 17 YDVLQKIMK---DDVAAFIQ--QESLTPKLNKNSISKFNQYTERDKFLG---DT 66
 QY 80 PNFVKGHVELYVNNRKIKIRYLQCPCKIKYDEREFRELITSHMQRTLNAXKIDYM 139
 DB 67 DYIAKQYKYLKNNHGVADVSEETPELTAEEAKIKRLKILINPASKKASLAQVUL 126
 QY 140 NHRDLV---AMAADICTAITNGEYQKGLYLYGPFGTGKSFILGAIANQLSKK-VRS 195

Db 127 DDLGRLPVFEKLAFAEQYPAIR-----KGLYVGFQGVKSPMAVALAHNDSEKRGVSS 181
QY 196 TTYLYDEEFTITLKGFGDSFEKKLRVREANILMDIDGAEVTPWVDEVTGPLLHTR 255
Db 182 TLHAFSPFYVDVNAISDGNVKTLYVDEIKLSEVLIIDDIGABOSTWVRDEIIQVLIQYR 241
QY 256 MYHELPTFPSSNDYSELHMLMTRDGEKKT-KAARITREVKSLSTPYELSGENFR 311
Db 242 MGENLPTFTSNFEDLEKHPAKVKGNDETWEARKVERIKRYLAEETRLLEGVNR 298

RESULT 14

ABBS3757
ID ABB53757 standard; Protein; 291 AA.

AC ABB53757;

DT 16-MAY-2002 (first entry)

DE Lactococcus lactis protein pili4.

KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

OS Lactococcus lactis IL1403.

PN FR2807446-A1.

PD 12-OCT-2001.

PF 11-APR-2000; 2000FR-0004630.

PR 11-APR-2000; 2000FR-0004630.

PA (INRG) INRA INST NAT RECH AGRONOMIQUE.

PI Bolocline A, Sorokline A, Renault P, Ehrlich SD;

DR WPI; 2002-043418/06.

PT New nucleotide sequence useful in the identification or Lactococcus

PS Claim 6; SEQ ID No 459; 2504bp; French.

CC The present invention relates to a Lactococcus lactis nucleotide

CC sequence (ABA90521) and related proteins (ABB53300-ABB5621). The

CC nucleic acid sequence is useful in the detection and/or amplification of

CC related species. The proteins of the invention are useful for the

CC biosynthesis or biodegradation of a composition of interest. The

CC production of yogurt and cheese.

CC Note: The sequence data for this patent is based on equivalent patent

CC WO20017734 (published 18-OCT-2001) which is available in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 291 AA;

QY Query Match 6.8%; Score 112; DB 23; Length 291;

Db Best Local Similarity 21.8%; Pred. No 0.088;

QY Matches 57; Conservative 36; Mismatches 94; Indels 74; Gaps 11;

Db 132 -----AKLDIYVNHDRDLVYMAADDICATITNGQVKGKYL 169

QY 88 SKYQKSLTKTOSVGRKSLMFAFNTFKVNGEBOVNLQAORINAEYIQQGRFNTVFV 147

QY 170 YGPFGRKSFILGAIANO-----LKSKKVRSTIYLPFIRTLKGFGDSFEKKLR--- 222

Db 148 -CGAGRGKSHLAMLIOEVNENIKDKFSTLFINISELIREIKSNWNYSDTAEERLTT 206
QY 223 -VREANILMDIDGAEV-----TPWRDDEVIGPLHYRVHLLPFFSSNPFYSEHNL 277
Db 207 LMRVDLAVLDLGTESTFSSKNSWQ-----GVYINYNARGNITITSLTGKEM---- 258
QY 278 AMTRDGEKTKAARITREVK 298
Db 259 ---RSSYDKIVSRIMEGSKN 276

RESULT 15

AAU38917
ID AAU38917 standard; Protein; 866 AA.

AC AAU38917;

DT 16-JAN-2002 (first entry)

DE C. pneumoniae CT113 homologue CPn0695.

KW Chlamydia; sexually transmitted disease; PID; antibacterial;

KW pelvic inflammatory disease; antigen; trachoma; gynecological;

KW acute respiratory tract infection; atherosclerosis; male infertility;

OS Chlamydia pneumoniae.

PN WO200181379-A2.

PD 01-NOV-2001.

PF 23-APR-2001; 2001WO-US13081.

PR 21-APR-2000; 2000US-198853P.

PR 20-JUL-2000; 2000US-219752P.

PA (CORI-) CORIXA CORP.

PI Bhatia A, Probst P, Stromberg EJ;

DR WPI; 2001-616771/71.

DR N-PSDB; AAS57049.

PT New polynucleotide for treating Chlamydia infections encodes a

PT polynucleotide containing an immunogenic portion of a Chlamydia

PS disclosure; Page 192-193; 208bp; English.

CC The invention relates to isolated polynucleotide encoding at least

CC a partial Chlamydia protein which is an antigenic fragment, or the

CC complements, fragments, homologues and variants, and antibodies

CC raised against the antigenic proteins (or fragments). The nucleic

CC acids, proteins and antibodies are used to diagnose and treat Chlamydia

CC infections (e.g. a sexually transmitted disease, pelvic inflammatory

CC disease (PID), acute respiratory tract infection, trachoma,

CC atherosclerosis and coronary heart disease) in a patient, and in

CC the treatment of male infertility. The compounds of the invention are

CC also useful for detecting the presence of Chlamydia in a patient, and

CC stimulating and/or expanding T cells specific for a Chlamydia protein.

XX The present sequence represents a Chlamydia antigen.

QY Query Match 6.7%; Score 110; DB 22; Length 866;

Db Best Local Similarity 22.0%; Pred. No 0.56;

QY Matches 81; Conservative 71; Mismatches 123; Indels 94; Gaps 20;

Db 422 IVQGEAKIKRGSPSQGEADAMQKSI-----DALREBLASRLGMDDEKKLISGL 471

```
QY 65 KDQKHYDGHKADCPNFVKGHVPELYVDNNRI-KIRYLOCPC---KIKYDE----- 112
Db 472 KEKNSLESMEFSE-----EBAERVADYRVVAELRYSILPQLEBEIKQDEASLNQRDN 524
QY 113 -----ERPEALI---TSHMQR--DTIAKLKDIYMHDRDL-----DVMAADDI 154
Db 525 RLQOEYDERLIAQVANNMTGI PVQKMLEGEAEKLLIEBSLEERVYVGOFPFAVSADSI 584
QY 155 CTA---ITNGEOVKGLYIY-GPFGTGKSPFIGAIAIANOI KSKVSTIILPEF-----IR 205
Db 585 RAARVGLANDPORPLGVFLFLGPTGVGKTELAKALADLLFNKEAMVRPDMSEYMEKHSIS 644
QY 206 TLKG-----GFKDGSPEKKLHVRREANIIMLDDIGAEVYTPWVRDEVIGPLHY----- 254
Db 645 KLIGSSPGYVGYEBSLSLEALRRRYPYSVVLPDEIERKAD-----KEVLNILLQVFDGDI 698
QY 255 -----RMVHELPTFF--SSNFDYSELHHLAMTRDGEKTKARITIERVKSLSPTY--- 303
Db 699 LTDGKKRRVNCCKNALFTMTSNIGSPELADY--CSKKGSELTK-----EAILSVVSPVLKR 751
QY 304 FLGGENFRN 312
Db 752 YLSPE-FMN 759
```

Search completed: December 15, 2003, 15:13:42
Job time : 47.6369 secs

APPLICANT: PhageTech, Inc.
 TITLE OF INVENTION: Compositions and methods involving an essential Staphylococcus Au
 TITLE OF INVENTION: gene and its encoded protein
 FILE REFERENCE: 21715/1000
 CURRENT APPLICATION NUMBER: US/09/470,512A
 CURRENT FILING DATE: 1999-12-12
 NUMBER OF SEQ ID NOS: 16
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 15
 LENGTH: 278
 TYPE: PRF
 ORGANISM: Bacillus halodurans
 US-09-470-512A-15

Query Match 34.6%; Score 568; DB 4; Length 278;
 Best Local Similarity 40.6%; Pred. No. 1.4e-48;
 Matches 117; Conservative 61; Mismatches 90; Indels 20; Gaps 7;

36 PVKQPLEAHRAILTNAMIDEDLVNLYQEKYKQKHYDGH-KPADQPNFVKGHVPELYD 93
 1 PNYQFLERH-PSLSPTTEGGLSKLYEYQKESHC-AHCPGLQKCPNLMKGYQPTLYE 58
 94 NNRIRKYQCPCKIKYDERFEALITSHMORDTNAKLDI-----YNNHRDL 146
 59 RUSLEISYSPCKPEERERKKKSLIRSLUIPKELIAKFDVSESPGRSIASHR-ALE 117
 147 VMAADICTALTNGQVGLYIGPFGKSFILGAINOLKSKKVRSTIYLPFIRIT 206
 118 FALSAP-----GDDGGLYLYGKFGVGTFLMGAINELMDRDISITIVVDPFEE 170
 207 LKGFQDGSFEKKLHVREANILMLDIGAEEYTPVWRDEVIGPLHYHVMHLPFFSS 266
 171 LKQSIDGTFQQLDVKQAQVILFDIGLETWTSVWRDVLGVIQYIMELGPLFIS 230
 267 NPDYSELHHLAMT-RDGEKTKARIIRKVSLSPPYFLSGENFENN 313
 231 NVDYDELEHIAVNDKSGTELLAKKVRMERIRHYTYSVWVGQNRREH 278

Db

RESULT 6
 US-09-107-532A-5580
 Sequence 5580, Application US/09107532A
 Patent No. 6583275
 GENERAL INFORMATION:
 APPLICANT: Lynn A Doucette-Stamm and David Bush
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 NUMBER OF SEQUENCES: 7310
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 STREET: 100 Beaver Street
 CITY: Waltham
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02354
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Ariniello, Pamela Deneka
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 5580:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 310 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (8) LOCATION 1...310
 SEQUENCE DESCRIPTION: SEQ ID NO: 5580:
 US-09-107-532A-5580

Query Match 27.3%; Score 449; DB 4; Length 310;
 Best Local Similarity 33.5%; Pred. No. 1.2e-36;
 Matches 106; Conservative 69; Mismatches 123; Indels 18; Gaps 7;

5 GSIMQFSTINTSQDFEKRIKIKYVINDVYKQFLFAHRAELTNAMIDEDLVNLYQ 64
 4 EDVGKELRII-SKNYQEKYKEMIAEVLNDADVQAFSAHQEBILQADIERKSYAKLYE 62
 65 KQQKHVGHKPADCPNPFVKGHVPELYDNNRIRKYQCPCKIKYDERFEALITSHH 124
 63 VQEKRYQKN--DPTMAIPGTEPKLTILFHYDVITYFTEALLARQOEIRSKYKAND 119
 125 MORDTNAKLDIYNNHRDL-----VMAADIC--TAINGEQVGLYLYGPGTK 177
 120 MKDQISAFENF-----ERTDGRGASLEALDIPFKYTAADPKGN-KGLYLYGSGICK 173
 178 SFILGAINOLKSKKVRSTIYLPFIRITLKGFGDGSFEKKLHVREANILMLDIGAE 237
 174 TYLGATRAELAIAGFTTLVHFPTFAVEMKQAIQDQVAEKLDVAKRSPILMIDIGAD 233
 238 EYTPVWRDEVIGPLHYHVMHLPFFSSNPFYSELEHHLAMTRDG-EKTKARIIRV 296
 234 AMSSWIRBVFVGLIYQRMQEQLPFFSSNFTMDELOHLSVYQRDREPLAKRIMERI 293
 297 KSLSTPYFLSGENFENN 312
 294 RYLTKEIENTGRNRN 309

Db

RESULT 7
 US-09-328-352-5963
 Sequence 5963, Application US/09328352
 Patent No. 6562958
 GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 5963
 LENGTH: 268
 TYPE: PRF
 ORGANISM: Acinetobacter baumannii
 US-09-328-352-5963

Query Match 7.5%; Score 124; DB 4; Length 268;
 Best Local Similarity 25.6%; Pred. No. 0.00024;
 Matches 34; Conservative 28; Mismatches 63; Indels 8; Gaps 4;

107 KIKYDERFEALITSHMORDTNAKLK-DIYNNHRDLVMAADICTAIT--NGNQ 163
 102 KYRBEPEVAVNTAVSKSKSQKTLNPQFTFSFVGRSN--QMAETGRKVLTOIGASQ 158
 164 VAGLVLYGPGTKSFILGAINOLKSKKVRSTIYLPFIRITLKGFGDGSFEKKH 221

Db 159 HNPFLYGPFGTGLTHMQAVGNALLQAKPNARVMTVSQSFQDVFSSLQXKGVBEFFK 218
Qy 222 RVEANILMDDI 234
Db 219 NCRSLDLLVDDI 231

RESULT 8

US-09-198-452A-163
Sequence 163, Application US/09198452A
Patent No. 6559294
GENERAL INFORMATION:
APPLICANT: Griffiths, R.
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
FILE REFERENCE: 9710-003-999
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 6849
SEQ ID NO 163
LENGTH: 872
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-198-452A-163

Query Match 6.7%; Score 110; DB 4; Length 872;
Best Local Similarity 22.0%; Pred. No. 0.034;
Matches 81; Conservative 71; Mismatches 123; Indels 94; Gaps 20;

Qy 7 IMQFKSIINTSGDFEKRIKIKKEVINDPVQFLAHRALTNAMI--DEDLNTQVEY 64
Db 428 IVKQEAIRKQSPSYQREADQMOKSI-----DALREELASRLGWDDEKGLISGL 477
Qy 65 KDQKHYDGHKADCPFVKGHVELYVNNRI-KIRYLOQPC--KIKYDE----- 112
Db 478 KKKKNSLESKMSF-----EEMERVADYNRAVELYSLIPQLEBEIKQDEASLNGRDN 530
Qy 113 -----ERPEALII--TSHMQR--DTLNAKLKDIYNNHRDL-----DYAMADDI 154
Db 531 RLQOEYDERLIQAVNMTGIPVQKMLEGEAEKLLLESLEBERRVGGPFAVSADSI 590
Qy 155 CTA---LTNGEYVGLVY--GPGTGSFLIGATNOLKSKKVVSTIYVPEF-----IR 205
Db 591 RAARVGLNDPQRPGLFGLPTGVGKTELAKALDLLENKEMARPDSEYMEKKSIS 650
Qy 206 TLKG-----GFKDGFEEKLHVRVREANILMDDIGAEVTPVVRDEVIGPLHY----- 254
Db 651 KLIGSSPGVGYEGEGSLSEALRRPYSVVLFDIEIKAD-----KEVLNILLQVDDGI 704
Qy 255 -----RMVHELPTFP--SSNFYSELEHILAMTRDGEKTKARITERYKSLSTPY--- 303
Db 705 LTDDKKRKVNCNALFTMTSNIGSPELADY--CSKKSSELTK-----EALISVSPVLKR 757
Qy 304 FLGSEFRN 312
Db 758 YLSPE-FRN 765

RESULT 9

US-08-938-105-3
Sequence 3, Application US/08938105
Patent No. 6353151
GENERAL INFORMATION:
APPLICANT: Leinwand, Leslie A.
TITLE OF INVENTION: TRANSGENIC MODEL FOR HEART FAILURE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln St., Suite 3500
CITY: Denver

STATE: CO
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938,105
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crook, Marnell M.
REGISTRATION NUMBER: 31,071
REFERENCE/DOCKET NUMBER: 3595-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1886 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
US-08-938-105-3

Query Match 6.4%; Score 105; DB 4; Length 1886;
Best Local Similarity 24.9%; Pred. No. 0.33;
Matches 58; Conservative 24; Mismatches 79; Indels 72; Gaps 10;

Qy 18 SDFEKRIEKIKKEVINDPVQFLAHRALTNAMIDEDLNTQVEYKDQKHDKH-- 75
Db 791 SATTEKEMAMKEPFGVDALEKSPARKLEKRV-----SLQKNDQLQVQAEQDN 846
Qy 76 PAD-----CPNFVKGHVELYVNNRIKIRYLOQPCRK-----YDEERPEALITSHM 125
Db 847 LADAERCDQLK-----NKI-----QLEAKVKEWTERLEDEEMNAALTLAKRK 891
Qy 126 QRTLNAKLKDIYNNHRDLVAMADDICTALTNEQVYGLVYCPFGKSFILGATA 165
Db 892 LDECESELKDI-----DDELTLAKVKEKH-----A 919
Qy 186 NOLSKKVRSTIYVPEFIRTLKGFQDGFEEKLHVRVREANILMDDIGAE 238
Db 920 TENKYNLTENMAGDEIITKL-----TKKK--ALQEAHQALDDLAEE 963

RESULT 10

US-09-134-001C-5153
Sequence 5153, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5153
LENGTH: 304
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5153

Query Match 6.2%; Score 102.5; DB 4; Length 304;
Best Local Similarity 21.4%; Pred. No. 0.04;
Matches 33; Conservative 35; Mismatches 59; Indels 27; Gaps 4;

Query Match 6.0%; Score 98; DB 2; Length 1964;
Best Local Similarity 20.7%; Pred. No. 1.8;
Matches 75; Conservative 53; Mismatches 110; Indels 124; Gaps 20;

QY 23 KRIEKT-----KEVINDDP-VKQFLHARAEITNMIDEDLVNLOEYKQOQK-----69
DB 1124 KNIKELLPYNNKSTIVKYGNLVKENSLLYQKELLSAVMMKDDQVITDIYNNKQTANKLL 1183
QY 70 HYDGHK-----FADCPNFKGHPVDELVDNNRIKIRYLQCPCKIKYBERFEAE 118
DB 1184 HYNDHSEKFDLKYQDTPANLPYNNLGNTGLYTPN-----QFLYRDSIVKE 1231
QY 119 LITSHHMQ-----RDTL-----NAKLDIYM-----NHRDL-----DVAMA 150
DB 1232 VLP--ELQKLDYQSDAIRKTLGISPEVKLTLEYLEDQFSKTKONLGSLSKLLSADAGLA 1289
QY 151 ADDICTAINTGEBOVKGLYLGPGTGKSPILGAIAN-----QLSKKYRSTIITLPER 203
DB 1290 SDN---SVTRG-----YLVDKIKNNKEALLGLTYLERWYNNFYQGVNKKDLVMYHPDF 1340
QY 204 -----IRTLKGFKDGSFEKKLHVRVANILM---LDDIGAEVTPWVRDEVI 248
DB 1341 FGKNTSPDLTLIELKSGFN-----NLLAKNNVDYTGISLASQHGATDLF 1386
QY 249 GPLHATRYMVELPTFFSSNPDYSELEHHLAMTRDGEKTKAARIIRVKSISTPYPLSGE 308
DB 1387 STLEHRYKVF-LPNTSNNDWPKSETKAYIV-----EKS---TIEBVK---TKQGLAGT 1433
QY 309 NF 310
DB 1434 KY 1435

RESULT 14
US-08-790-912-2
Sequence 2, Application US/08790912
Patent No. 5976542

GENERAL INFORMATION:

APPLICANT: Weiser, Jeffrey N.

APPLICANT: Plant, Andrew G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT
OF INFECTIONS OF STREPTOCOCCUS PNEUMONIAE INFECTION

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: PAINTCH SCHWARZE JACOBS & NADEL, P.C.

STREET: 1601 Market Street, 36th Floor

CITY: Philadelphia

STATE: Pennsylvania

COUNTRY: USA

ZIP: 19103-2398

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/790,912

FILING DATE: 29-JAN-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/026,939

FILING DATE: 23-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Leary, Kathryn

REGISTRATION NUMBER: 36,317

REFERENCE/DOCKET NUMBER: 7600-401

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 567-2020

TELEFAX: (215) 567-2991

TELEX: 831-894

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2052 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-790-912-2

Query Match 6.0%; Score 98; DB 2; Length 2052;
Best Local Similarity 20.7%; Pred. No. 1.9;
Matches 75; Conservative 53; Mismatches 110; Indels 124; Gaps 20;

QY 23 KRIEKT-----KEVINDDP-VKQFLHARAEITNMIDEDLVNLOEYKQOQK-----69
DB 1197 KNIKELLPYNNKSTIVKYGNLVKENSLLYQKELLSAVMMKDDQVITDIYNNKQTANKLL 1256
QY 70 HYDGHK-----FADCPNFKGHPVDELVDNNRIKIRYLQCPCKIKYBERFEAE 118
DB 1257 HYNDHSEKFDLKYQDTPANLPYNNLGNTGLYTPN-----QFLYRDSIVKE 1304
QY 119 LITSHHMQ-----RDTL-----NAKLDIYM-----NHRDL-----DVAMA 150
DB 1305 VLP--ELQKLDYQSDAIRKTLGISPEVKLTLEYLEDQFSKTKONLGSLSKLLSADAGLA 1362
QY 151 ADDICTAINTGEBOVKGLYLGPGTGKSPILGAIAN-----QLSKKYRSTIITLPER 203
DB 1363 SDN---SVTRG-----YLVDKIKNNKEALLGLTYLERWYNNFYQGVNKKDLVMYHPDF 1413
QY 204 -----IRTLKGFKDGSFEKKLHVRVANILM---LDDIGAEVTPWVRDEVI 248
DB 1414 FGKNTSPDLTLIELKSGFN-----NLLAKNNVDYTGISLASQHGATDLF 1459
QY 249 GPLHATRYMVELPTFFSSNPDYSELEHHLAMTRDGEKTKAARIIRVKSISTPYPLSGE 308
DB 1460 STLEHRYKVF-LPNTSNNDWPKSETKAYIV-----EKS---TIEBVK---TKQGLAGT 1506
QY 309 NF 310
DB 1507 KY 1508

RESULT 15
US-08-820-170A-13
Sequence 13, Application US/08820170A
Patent No. 5831058

GENERAL INFORMATION:

APPLICANT: Teutonu, FUJIWARA

APPLICANT: Takeishi, WATANABE

APPLICANT: Masato, HORIE

TITLE OF INVENTION: HUMAN GENE

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sughrue, Mwon, Zimm, Macpeak & Seas

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: United States

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/820,170A

FILING DATE:

CLASSIFICATION: 536

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 293-7060

TELEFAX: (202) 293-7860

TELEX: 6491103

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 389 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-820-170A-13

Query Match 5.9%; Score 97; DB 2; Length 389;
Best Local Similarity 23.9%; Pred. No. 0.2;
Matches 62; Conservative 39; Mismatches 86; Indels 72; Gaps 14;

QY	19	QDFEKRIEKIKKAVINDPVKQFLFAHRAELTN--AMIDEDIANVLOE---YKDOQKHVD 72
DB	10	QDYRKRLLE-HKEI--DGRLEKELREQLK-ELTKQYKESNDLKALQSVGQIVGEVLKQULT 65
QY	73	GHKF---ADCPNFVKGHVPELVVDNNRIKIRYLQCPCKIKYDEERFEALITSHHMQRD 128
DB	66	EEKFIVKATNGPRYVVG-----CRQLDKSKLKPGTRVALDPTTL 105
QY	129	TINAKLKD-----LY-MNHRDLRLDVAMA-----ADDICTAITNGE----- 162
DB	106	TIKRYLPRREVDPLVYMNMSHEDPGNVSYSEIGLSQIRLEIREVIELPLTNPELFGVGI 165
QY	163	QVKGILYGPFGTGKSFILGAINQKLS---KKVSTII--YIPERITLKGGFQDGSFE 217
DB	166	PPKGCILYGPFGTKTLAAVAASQDCNFKVSSSIYDKYIGESARLIREMF----- 219
QY	218	KKLHVRVREANITLMDDIGA 236
DB	220	-NYARDHPCCILFMDRID 237

Search completed: December 15, 2003, 15:17:39
Job time : 18.3604 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:13:51 / Search time 30.5681 Seconds
(without alignments)
1904.368 Million cell updates/sec

Title: US-09-689-952-2

Perfect score: 1643
Sequence: 1 MCGGQSIMKQFKSIINTSQD.....ERVKSLSTPYFLSGENFRNN 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	110	6.7	866	12	US-09-841-260-113
2	110	6.7	866	14	US-10-007-693-113
3	105	6.4	248	12	US-10-287-274-435
4	102	6.2	389	15	US-10-128-714-3369
5	102	6.2	389	15	US-10-128-714-3369
6	101.5	6.2	827	12	US-10-032-585-7621
7	101	6.1	451	15	US-10-128-714-8421
8	100	6.1	653	15	US-10-156-761-11850
9	100	6.1	740	12	US-10-094-749-2991
10	99.5	6.1	1177	15	US-10-128-714-3421
11	98	6.0	393	15	US-10-128-714-3421
12	98	6.0	467	9	US-09-815-242-10395
13	97.5	5.9	867	12	US-09-841-260-109
14	97.5	5.9	867	14	US-10-007-693-109
15	97.5	5.9	1179	15	US-10-128-714-8493

16	97	5.9	389	10	US-09-976-165-13	Sequence 13, App1
17	97	5.9	389	12	US-10-205-219-165	Sequence 165, App
18	97	5.9	389	12	US-10-342-276-13	Sequence 13, App1
19	96.5	5.9	442	9	US-09-815-242-13552	Sequence 13552, A
20	96.5	5.9	652	9	US-09-351-794A-2	Sequence 2, App1
21	96.5	5.9	682	12	US-10-032-585-7047	Sequence 7047, Ap
22	96	5.8	454	9	US-09-815-242-11167	Sequence 11167, A
23	96	5.8	1368	12	US-10-032-585-7544	Sequence 7544, Ap
24	95	5.8	445	9	US-09-815-242-5010	Sequence 5010, Ap
25	95	5.8	447	9	US-09-815-242-10908	Sequence 10908, A
26	94.5	5.8	411	12	US-10-032-585-7284	Sequence 7284, Ap
27	94.5	5.8	796	12	US-10-032-585-7201	Sequence 7201, Ap
28	94.5	5.8	942	12	US-10-334-703-51	Sequence 51, App1
29	94.5	5.8	942	12	US-10-396-122-51	Sequence 51, App1
30	93.5	5.7	577	9	US-09-815-242-10193	Sequence 10193, A
31	93.5	5.7	1576	12	US-10-396-122-81	Sequence 91, App1
32	93.5	5.7	1940	12	US-09-738-630-99	Sequence 99, App1
33	93	5.7	402	9	US-09-925-301-984	Sequence 984, App
34	92.5	5.6	574	10	US-09-764-868-774	Sequence 774, App
35	92	5.6	870	15	US-10-156-761-12048	Sequence 12048, A
36	91.5	5.6	284	9	US-09-861-451A-20	Sequence 20, App1
37	91.5	5.6	350	10	US-09-738-626-6009	Sequence 6009, Ap
38	91.5	5.6	401	12	US-10-032-585-7325	Sequence 7325, Ap
39	91.5	5.6	1240	12	US-10-032-585-7319	Sequence 7319, Ap
40	91	5.5	2125	10	US-09-919-172-29	Sequence 29, App1
41	91	5.5	2649	12	US-10-205-219-169	Sequence 169, App
42	91	5.5	2649	12	US-10-341-434-220	Sequence 220, App
43	91	5.5	2649	12	US-10-341-434-230	Sequence 230, App
44	90.5	5.5	304	9	US-09-864-761-43043	Sequence 43043, A
45	90.5	5.5	508	11	US-09-769-787-102	Sequence 102, App

ALIGNMENTS

RESULT 1
US-09-841-260-113
Sequence 113, Application US/09841260
Publication No. US20030175700A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Probst, Peter
APPLICANT: Stromberg, Erika Jean
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
FILE REFERENCE: 210121.515
CURRENT APPLICATION NUMBER: US/09/841,260
CURRENT FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 140
SEQ ID NO 113
LENGTH: 866
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-841-260-113

Query Match	6.7%	Score 110,	DB 12,	Length 866;
Best Local Similarity	22.0%	Pred. No. 0.12,		
Matches	81,	Conservative	71,	Mismatches 123, Indels 94, Gaps 20;
QY	7	IMKQFKSIINTSQDEKRIKKEVINDDPVKQFLKRAHRLTNMI--DDLNLVLORY	64	
DB	422	IVKQAIKREQSPSYQEBADMQKSI-----DLRELASLRIGMDKKLSL	471	
QY	65	KDQQRHVDGHRFADCPNFVKGVPELVYDNNRI-KIRYQCC--KIKYD-----	112	
DB	472	KEKKNSLSMKFSF-----EAEARVADYNNVALRYSLIPQLREIKQDASINQRDN	524	
QY	113	-----ERFRELII--TSNHR--DTLNKIKDIYNNHRDL-----DYMAADI	154	
DB	525	RLQGEVDERLIAQVAVNWGTIPVQKMLGEAEKLLIIESSLEKVVGGPPFAVSASDSI	584	
QY	155	CTA---ITNGEVDKLVLY-GPFGTGSFTILGAINOLKSKKVRSTIYLPF-----IR	205	

```

Db      585 RAARVGLNDPQRLVFLGPTGKGTBLAKALADLLFNKEBAMVAFDMSVEYKKSIS 644
Qy      206 TLKG-----GPKGSEPKLHREANILMLDIGAEBVTWVRDVIQPLHY----- 254
Db      645 KLIGSSPGYGYEGGSLSEALRRPYSVVLFDIEKAD-----KEVLNILLQVFDGI 698
Qy      255 -----RMVHELPTFF--SSNFDYSELEHHLAMTRDGEKTKAARIERYKSLSTPY--- 303
Db      699 LTGDKRKKVNCNKLFTMTSNIGSPELADY--CSKKSSELTG-----EALISVSPVLKR 751
Qy      304 FLGGENFRN 312
Db      752 YLSPE-FMN 759

```

RESULT 2

```

US-10-007-693-113
; Sequence 113, Application US/10007693
; Publication No. US2002014676A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Probec, Peter
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 113
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-007-693-113

```

Query Match 6.7%; Score 110; DB 14; Length 866;

Best Local Similarity 22.0%; Pred. No. 0.12;

Matches 81; Conservative 71; Mismatches 123; Indels 94; Gaps 20;

```

Qy      7 IMQPKSIINTSDPEKRIKIKKEVINDDPVQPLFNAHAEITNMI--DEDLNVQEV 64
Db      422 IVKQEAIRKQSPSYQSEADAMQKSI-----DALREELASLRLGWDDEKQLISGL 471
Qy      65 KDOQKHVDGKHFADCPFVKGHVLYVNNRI-KIRYLOQPC--KIKYDE----- 112
Db      472 KEKKNLSKMKFS-----EEHRYADYNRAVELATYSLLPOLBEITKQDEHSLNGRDN 524
Qy      113 -----ERPEAEI--TSHMQR--DTLNALKDQIYNNHRDL-----DYMAADDI 154
Db      525 RLQGEVDERLIAQVAVNMGTGIPQKMLBGAELKLLLEESLEBRYVVGOPFAVSAVDSI 584
Qy      155 CTA---ITNGEQVGLVLY-GPFGTGSFLIGALANOLKSKVSTIILPEF-----IR 205
Db      585 RAARVGLNDPQRLVFLGPTGKGTBLAKALADLLFNKEBAMVAFDMSVEYKKSIS 644
Qy      206 TLKG-----GPKGSEPKLHREANILMLDIGAEBVTWVRDVIQPLHY----- 254
Db      645 KLIGSSPGYGYEGGSLSEALRRPYSVVLFDIEKAD-----KEVLNILLQVFDGI 698
Qy      255 -----RMVHELPTFF--SSNFDYSELEHHLAMTRDGEKTKAARIERYKSLSTPY--- 303
Db      699 LTGDKRKKVNCNKLFTMTSNIGSPELADY--CSKKSSELTG-----EALISVSPVLKR 751
Qy      304 FLGGENFRN 312
Db      752 YLSPE-FMN 759

```

RESULT 3

```

US-10-287-274-435
; Sequence 435, Application US/10287274
; Publication No. US20030181408A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn

```

```

; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyekind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THEREOF
; FILE REFERENCE: ELITTA.008DV1
; CURRENT APPLICATION NUMBER: US/10/287,274
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 435
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-435

```

Query Match 6.4%; Score 105; DB 12; Length 248;

Best Local Similarity 26.7%; Pred. No. 0.062;

Matches 43; Conservative 30; Mismatches 66; Indels 22; Gaps 7;

```

Qy      140 NHRDRLDVANA--ADICTAINGQVKGLYYGFPGTGSFLIGALANQLSKVYST 196
Db      80 NEGQVALSQAKSIDELMTGCTN-----FAFGKPGTKMHAIAIGRLKDGQTVI 133
Qy      197 IYLPERIRTKGFPDG-SFEKGLRVRANILMLDIGAEBVTWVRDVIQPLHYR 255
Db      134 VVTVAOVMSALHASYDGGSGEFLRELCVDLVLVDGIGQET--KNEQV--VLH-Q 187
Qy      256 MVHELPTFFSSNFDYSELEHHLAMTRDGEKTKAARIERY 296
Db      188 IYDRRTASMSVGMLTNLTNANEAANKTILGE-----RIMDRM 222

```

RESULT 4

```

US-10-128-714-3369
; Sequence 3369, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengdi
; APPLICANT: Fishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Broshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3369
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3369

```

```

US-10-128-714-3369
; Sequence 3369, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengdi
; APPLICANT: Fishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Broshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3369
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3369

```

Query Match 6.2%; Score 102; DB 15; Length 389;

Best Local Similarity 21.5%; Pred. No. 0.23;

Matches 67; Conservative 60; Mismatches 104; Indels 80; Gaps 15;

```

QY 21 FEKRIKIKKEVINDDPVKQFLBAHRAELTN--AMIDEDINVLQOEYKQOKHYDGH---K 75
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 7 YRKIKESMKLEIIQGAVALRLERAGRDVNSRVRLRLREELGLQ-----QGGSYGGEVVKV 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 76 FADCPNFVKGHVELYVDN--NRIKIRYLOCPCKIKYDEERFEAE-----LITSHH 124
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 MSTRKVLVVKHPEKGYVVDIADGVDTLTKLVGKRVALLSDSYLKEKMLPSSVDPVLVSLMM 122
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 125 MOR-----DTLNAKLDIYMNHRDLVMAADDICTALTNGEQVGLVXGPFGT 175
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 123 VEKVPDSTYDMIGLDOQIKELI-----KEVTELGKHPLEBSJ-GIAQPGVLLYGPFGT 177
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 GKSPIIGAIANOLSKSKVR---STII--YLPEFIRTLKGGFKDGSFEKTLHVRREANIIM 230
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 GKTLLARAVAHHTDCKRFIRVSGSELVQYIGBSRWRLPF-----VWAREHAPSIIIF 230
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 LDDIGAEVTPWVRDEVIGPLHYRMVHELPTFPSSNFDY-----SELEH--HLAMTRD 282
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 MDEI-----DSIG-----SSRIDSGSGDSGVQRTMELANQD 264
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 GEEKTKAARI 293
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 265 GEPTPKIKIT 275
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 5

```

US-10-128-714-8369
; Sequence 8369, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Weng
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128, 714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8369
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8369

```

Query Match 6.2%; Score 102; DB 15; Length 389;
 Best Local Similarity 21.5%; Pred. No. 0.23;
 Matches 67; Conservative 60; Mismatches 104; Indels 80; Gaps 15;

```

QY 21 FEKRIKIKKEVINDDPVKQFLBAHRAELTN--AMIDEDINVLQOEYKQOKHYDGH---K 75
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 7 YRKIKESMKLEIIQGAVALRLERAGRDVNSRVRLRLREELGLQ-----QGGSYGGEVVKV 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 76 FADCPNFVKGHVELYVDN--NRIKIRYLOCPCKIKYDEERFEAE-----LITSHH 124
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 MSTRKVLVVKHPEKGYVVDIADGVDTLTKLVGKRVALLSDSYLKEKMLPSSVDPVLVSLMM 122
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 125 MOR-----DTLNAKLDIYMNHRDLVMAADDICTALTNGEQVGLVXGPFGT 175
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 123 VEKVPDSTYDMIGLDOQIKELI-----KEVTELGKHPLEBSJ-GIAQPGVLLYGPFGT 177
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 GKSPIIGAIANOLSKSKVR---STII--YLPEFIRTLKGGFKDGSFEKTLHVRREANIIM 230
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 178 GKTLLARAVAHHTDCKRFIRVSGSELVQYIGBSRWRLPF-----VWAREHAPSIIIF 230
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 231 LDDIGAEVTPWVRDEVIGPLHYRMVHELPTFPSSNFDY-----SELEH--HLAMTRD 282
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 MDEI-----DSIG-----SSRIDSGSGDSGVQRTMELANQD 264
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 GEEKTKAARI 293
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 265 GEPTPKIKIT 275
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 6

```

US-10-032-585-7621
; Sequence 7621, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032,585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7621
; LENGTH: 827
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7621

```

Query Match 6.2%; Score 101.5; DB 12; Length 827;
 Best Local Similarity 24.2%; Pred. No. 0.74;
 Matches 71; Conservative 41; Mismatches 96; Indels 85; Gaps 17;

```

QY 19 ODFE-KRIEK--IKK-----EVINDPVKQFLBAHRAELTNAMIDED-----LNVYL 61
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 64 KDFOLKRMKKNVLEKSIASLKVIRQDDEBERF-----GDRNGIDGSESEANDRECENDL 118
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 QEYKD-----QKHYDGHKFPDCEPVKCH---VELYVDNR-----IKIRY 101
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 MEYKDSNAINKSVSLMNQKTDGE--SDGVNNEBQQAQDTTELKGNRRKAKGSAKTOL 176
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 102 LQCPCKIKYDEERFE-----AELITSHMORDTLNAKLDIYMNHRDLVMAADDIC 155
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 177 KQQRKIDYSTNIDLSISGVESVTTQLLEIILGLPILHPEIYSS----- 221
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 156 TAITNGEQVGLVXGPFGTKSFIIGAIANOLSKSKVSTIIYLPFIRTLKGGFKDGS 215
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 222 ---TGVEPRGVLLVGPFCGKTTIANMALGELKVPFIN---ISAPSVSGM-----SGE 270
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 216 FEKGLHVR-REAN-----TLMDDIGAEVTPWVRDEVIGPLHYRMVHELPT 262
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 271 SEKKLREIFEEAKQIAPCLIFMDEIDA--ITP-KRDGAQRMEKRIYVAQLLT 320
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7

```

US-10-128-714-8421
; Sequence 8421, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and

```

TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8421
LENGTH: 451
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-8421

Query Match 6.1%; Score 101; DB 15; Length 451;

Best Local Similarity 20.8%; Pred. No. 0.35;

Matches 52; Conservative 51; Mismatches 87; Indels 60; Gaps 12;

QY 34 NDDPVKQFLAHRAELTNMAIDDE--LNVLOEYKDOQKHYD---GHKPADCPKPVYGHV 87
DB 63 SDBREBQALDYKSKLLELEWEKAKLNLWGKIDQREPDISSENTKALQSVQIIGEV 122
QY 88 PELVYDNNRIKIRYLQCP---CKIKYDERFE-----AELITSHHMDRT----- 129
DB 123 LK-QLBBERIVYVASSSPRRVVGGRSKVDKSKQGRVALDMTTLTLMMLPREVDPLV 181
QY 130 -----LNAKLDIYMHNRDLVMAADICTALTNGEYQVGLVLYG 171
DB 182 YNMSLEDPQGISFAGISGLNDQIREL---REVLELPLKPELPQRY-GIKPPGVLLYG 236
QY 172 PRTGKSFILGALANOLKS---KKVRSITL--YIPETIRLKGFKGSEKULHREA 226
DB 237 PRTGKTLARAAVASMETFLKVSATVYKYGESARLIREMF---GYAKE---HEP 289
QY 227 NILLMDIGA 236
DB 290 CILFMEIDA 299

RESULT 8

US-10-156-761-11850

Sequence 11850, Application US/10156761

Publication NO. US20030119018A1
GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHITO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11850
LENGTH: 653
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-11850

Query Match 6.1%; Score 100; DB 15; Length 653;

Best Local Similarity 39.2%; Pred. No. 0.74; Mismatches 20; Indels 10; Gaps 4;

Matches 29; Conservative 15; Mismatches 20; Indels 10; Gaps 4;

QY 167 LVLVGPFGTKSPILGALANOLKS---KKVR--STIYLPPIRTLKGFQKGSFEKUL 220
DB 350 LFIYSSGKTKLHLAIGHVYASLPGTRVRYVSSSEPFNFINISIRDS-KDSFPRK-- 406
QY 221 HVRVEMITLMDI 234
DB 407 -RYREMDIILVVDI 419

RESULT 9

US-10-094-749-2991

Sequence 2991, Application US/10094749

Publication NO. US20030219741A1

GENERAL INFORMATION:

APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YUKI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOTYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 08435/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2991
LENGTH: 740
TYPE: PRT
ORGANISM: Homo sapiens
US-10-094-749-2991

Query Match 6.1%; Score 100; DB 12; Length 740;

Best Local Similarity 21.3%; Pred. No. 0.88; Mismatches 97; Indels 78; Gaps 14;

Matches 61; Conservative 50; Mismatches 97; Indels 78; Gaps 14;

QY 5 QGIMKQPKSITINSQPFRIKIKKEVI--NDPVKQPLAARA---ELTNAMIBDUN 59
DB 212 KNVMKTFREEL--VNIKAFEVERQELASNCKKKEQALQAINAKELETLN-----N 261
QY 60 VLQEVYDQKHVDGHAFADCPNFVKGHVPELVYDNNRIKIRYLQCPCKIKYDERFEAL 119
DB 262 RMKKVDYERKQLRQRIIMDCBEY-----NMKI-----KLEQDVQILIEQL 302
QY 120 ITSHHQDPTLNAKLDIYMHNRDL---YMAADICTALTNGEYQVGLVYGPFGT 175
DB 303 -----QQR-----KATYQLONEKLEYNLQVTKRDESTVYKSGQKKRYNLHD--- 346
QY 176 GKSFILGALANOLKSKRVSTIYLPPIRTLKGFQKGSFEKULHVRVANITLMDIG 235
DB 347 -----ILNLRSKYAKQ--IKQFQENQSLTSYKRLVMQFKELQKAMRHFALIDDEK 397

QY 236 ABEVTPVDEVIPLHYRMVHELPFSSNPDYSEL--EHLIAM 279
 Db 398 FNEI-WLWNE-----BEADLIRAPADVRIHTHGL 430

RESULT 10

US-10-128-714-3493
 ; Sequence 3493, Application US/10128714
 ; Publication No. US20030119013A1

GENERAL INFORMATION:

APPLICANT: Jlang, Bo
 APPLICANT: Hu, Wengdi
 APPLICANT: Tishkoff, Daniel
 APPLICANT: Zamudio, Carlos
 APPLICANT: Eroshkin, Alexey M
 APPLICANT: Lemieux, Sebastien M
 TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
 FILE REFERENCE: 10182-018-999
 CURRENT APPLICATION NUMBER: US/10/128,714
 CURRENT FILING DATE: 2002-04-23
 PRIOR APPLICATION NUMBER: US 60/285,697
 PRIOR FILING DATE: 2001-04-23
 PRIOR APPLICATION NUMBER: US 60/287,066
 PRIOR FILING DATE: 2001-04-27
 PRIOR APPLICATION NUMBER: US 60/295,890
 PRIOR FILING DATE: 2001-06-05
 PRIOR APPLICATION NUMBER: US 60/303,899
 PRIOR FILING DATE: 2001-07-09
 PRIOR APPLICATION NUMBER: US 60/316,362
 PRIOR FILING DATE: 2001-08-31
 NUMBER OF SEQ ID NOS: 8603
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 3493
 LENGTH: 1177
 TYPE: PRF
 ORGANISM: Aspergillus fumigatus
 US-10-128-714-3493

Query Match 6.1%; Score 99.5; DB 15; Length 1177;
 Best Local Similarity 25.5%; Pred. No. 1.9;
 Matches 64; Conservative 33; Mismatches 89; Indels 65; Gaps 13;

QY 13 SIINTSODEBKREIKKEVINPDVKQF-----LEHRAELTNMIDELANTLOEY 64
 Db 742 SIIQAVEEMKSNTEQLKQD-ISDAKTRQAEAMKDIRIEKMSSEFNKND-DSKLAELQAT 799
 QY 65 KQOQKHVDGKHPADCFNFGVPELYVDNNRIKIRYLQCPCKIKYDEERF-EAEILITSH 123
 Db 800 LBSLK-----KCLAKNSNSVKALQKELQIS-----RLSEQVGSLSAAEQYAEAE----- 846
 QY 124 HMQRDTLNLAKLDIYNNHNRDLVMAAADICTAITNGEQVKLYVGPFGTKSPFLGA 183
 Db 847 -----NLIKQMEBIEQMKREO---ARDADHIAQALDEDEAK-----LTFDELEBE 891
 QY 184 IANQLSKKVRSTIITLPEFIRLTKGFGDGSFEKGL-HRVRAANTIMDDIGA----- 236
 Db 892 LEBALSKNSRLT-----BERGLEMKLGHQLEK-----LQKQQAATAQVAH 933
 QY 237 -BEVTPVDE 246
 Db 934 MEERHEMIAD 944

RESULT 11

US-10-128-714-3421
 ; Sequence 3421, Application US/10128714
 ; Publication No. US20030119013A1

GENERAL INFORMATION:

APPLICANT: Jlang, Bo
 APPLICANT: Hu, Wengdi
 APPLICANT: Tishkoff, Daniel
 APPLICANT: Zamudio, Carlos

APPLICANT: Eroshkin, Alexey M
 APPLICANT: Lemieux, Sebastien M
 TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
 FILE REFERENCE: 10182-018-999
 CURRENT APPLICATION NUMBER: US/10/128,714
 CURRENT FILING DATE: 2002-04-23
 PRIOR APPLICATION NUMBER: US 60/285,697
 PRIOR FILING DATE: 2001-04-23
 PRIOR APPLICATION NUMBER: US 60/287,066
 PRIOR FILING DATE: 2001-04-27
 PRIOR APPLICATION NUMBER: US 60/295,890
 PRIOR FILING DATE: 2001-06-05
 PRIOR APPLICATION NUMBER: US 60/303,899
 PRIOR FILING DATE: 2001-07-09
 PRIOR APPLICATION NUMBER: US 60/316,362
 PRIOR FILING DATE: 2001-08-31
 NUMBER OF SEQ ID NOS: 8603
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 3421
 LENGTH: 393
 TYPE: PRF
 ORGANISM: Aspergillus fumigatus
 US-10-128-714-3421

Query Match 6.0%; Score 98; DB 15; Length 393;
 Best Local Similarity 20.8%; Pred. No. 0.56;
 Matches 52; Conservative 50; Mismatches 88; Indels 60; Gaps 12;

QY 34 NDPVKQFLHRAELTNMIDED--LNLQYKQOQKHVD-----GHKPADCFNFGVHV 87
 Db 5 SDPERQALDEDKSLBLERWEALKNLRMGIKOLQRFDISEENIKALQSVGIGLV 64
 QY 88 PELVYDNNRIKIRYLQCP-----CKIKYDEERF-----ABLITSHMQRDT----- 129
 Db 65 LK-QLDEERFIVKASGPRVYVGCSSKVKDKSLKQGTVALDMTTLITIRMLPREVDPLV 123
 QY 130 -----LNALKDIIYNNHNRDLVMAAADICTAITNGEQVKLYLVG 171
 Db 124 YNMSLEDPEQISFAGIGLNDQIREL---REVILPLKNPFLQRV-GIKPPKGVLLYG 178
 QY 172 PRTGKSPFLGAIANOLKS---KVRSTII--YLPEFIRLTKGFGDGSFEKGLHVRBA 226
 Db 179 PRGTGKTLARAVASMEINPLKAVVSAIVDKIGESARLIEMF---GYAE-----HEP 231
 QY 227 NIMLDDIGA 236
 Db 232 CIIFMEIDA 241

RESULT 12

US-09-815-242-10395
 ; Sequence 10395, Application US/09815242
 ; Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari L.
 APPLICANT: Zyskind, Judith W.
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John D.
 APPLICANT: Carr, Grant J.
 APPLICANT: Yamamoto, Robert T.
 APPLICANT: Xu, H. Howard
 TITLE OF INVENTION: Identification of Essential Genes in
 FILE REFERENCE: ELITRA.011A
 CURRENT APPLICATION NUMBER: US/09/815,242
 CURRENT FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10395
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10395

Query Match
Best Local Similarity 27.1%; Pred. No. 0.72;
Matches 19; Conservative 21; Mismatches 28; Indels 2; Gaps 1;

167 LVIYGPRTGKSPFLGAIANOLSKKVRSTIYL--PEFRTLKGFQKDSFEKGLHVR 224
168 LFLYGGTGLCKTHLHVGNGIMAKPKNAVYVHSERFVQDVKALQNNNAIEEFKRYR 227
225 EANIMLDDI 234
228 SYDALIIDDI 237

RESULT 13
US-09-841-260-109
; Sequence 109, Application US/09841260
; Publication No. US20030175700A1
; GENERAL INFORMATION:
; APPLICANT: Bhadia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND DIAGNOSIS
; TITLE OF INVENTION: OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515
; CURRENT APPLICATION NUMBER: US/09/841,260
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 140
; SEQ ID NO 109
; LENGTH: 867
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-09-841-260-109

Query Match
Best Local Similarity 19.9%; Pred. No. 1.9; Length 867;
Matches 71; Conservative 71; Mismatches 110; Indels 105; Gaps 18;

18 SODEKRIEIKIKKEVINDPVKQFLNRAELTNAMIDEDLVLOEQKQKHGKFA 77
440 AEDMOKAIDRVKE-----LAALRLRW-----DEKGLITGKEKKNLNFKFA 484
78 DCNPFVKGAVPELYVNNRI-KIRYLQPC-----KITYDE--ERFEA 117
485 E-----BEAERTADNRVAELRYSLIPSEEBIHLEBALNQRDGRLLQEBVEDRLIA 537
118 ELI---TSHMOR--DTLNALKDVIYNNHRDL-----DVAMADDICTA---ITNGEOV 164
538 QVAMNTGIVQVQKMESEKLVLESLEERVVGQPPAIAVSDSIRARVGLSDPQP 597
165 KGLVLY-GPFGTGSFLGAIANOLSKKVRSTIYLPF-----IRTLKG-----GFK 212
598 LGVFLFGPTGVGKTELAKALABELFNKEBAMIRFDMTEWMEKGSVSKLGSPGYGYE 657
213 DGSFEKKLHVRANIMLDDIGAEVTPVNRDEVIGPLIHY-----RMHELP 261
658 EGGSLSEALRRRYSVVLPEIEKAD-----KEVFNILLQIPDDGILTDSKRRKVCN 711
```

```

262 TFE--SSNFDYSELHHLAMTRDSEKTKAARIER-----VKSISTEYPLS 306
712 ALPMTSNIGSOBLADYC-----TKKGTIVDKAVLSVAPALKNYFSSEFIN 759

RESULT 14
US-10-007-693-109
; Sequence 109, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
; APPLICANT: Bhadia, Ajay
; APPLICANT: Probst, Peter
; APPLICANT: Stromberg, Erika Jean
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.515C2
; CURRENT APPLICATION NUMBER: US/10/007,693
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 109
; LENGTH: 867
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-109

Query Match
Best Local Similarity 19.9%; Pred. No. 1.9; Length 867;
Matches 71; Conservative 71; Mismatches 110; Indels 105; Gaps 18;

18 SODEKRIEIKIKKEVINDPVKQFLNRAELTNAMIDEDLVLOEQKQKHGKFA 77
440 AEDMOKAIDRVKE-----LAALRLRW-----DEKGLITGKEKKNLNFKFA 484
78 DCNPFVKGAVPELYVNNRI-KIRYLQPC-----KITYDE--ERFEA 117
485 E-----BEAERTADNRVAELRYSLIPSEEBIHLEBALNQRDGRLLQEBVEDRLIA 537
118 ELI---TSHMOR--DTLNALKDVIYNNHRDL-----DVAMADDICTA---ITNGEOV 164
538 QVAMNTGIVQVQKMESEKLVLESLEERVVGQPPAIAVSDSIRARVGLSDPQP 597
165 KGLVLY-GPFGTGSFLGAIANOLSKKVRSTIYLPF-----IRTLKG-----GFK 212
598 LGVFLFGPTGVGKTELAKALABELFNKEBAMIRFDMTEWMEKGSVSKLGSPGYGYE 657
213 DGSFEKKLHVRANIMLDDIGAEVTPVNRDEVIGPLIHY-----RMHELP 261
658 EGGSLSEALRRRYSVVLPEIEKAD-----KEVFNILLQIPDDGILTDSKRRKVCN 711

262 TFE--SSNFDYSELHHLAMTRDSEKTKAARIER-----VKSISTEYPLS 306
712 ALPMTSNIGSOBLADYC-----TKKGTIVDKAVLSVAPALKNYFSSEFIN 759

RESULT 15
US-10-128-714-8493
; Sequence 8493, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengli
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Broshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
```

;; PRIOR FILING DATE: 2001-04-27
;; PRIOR APPLICATION NUMBER: US 60/235,890
;; PRIOR FILING DATE: 2001-06-05
;; PRIOR APPLICATION NUMBER: US 60/303,899
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: US 60/316,362
;; PRIOR FILING DATE: 2001-08-31
;; NUMBER OF SEQ ID NOS: 8603
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 8493
;; LENGTH: 1179
;; TYPE: PRT
;; ORGANISM: Aspergillus fumigatus
US-10-128-714-8493

Query Match 5.9%; Score 97.5; DB 15; Length 1179;
Best Local Similarity 25.5%; Pred. No. 3;
Matches 64; Conservative 33; Mismatches 91; Indels 63; Gaps 13;

QY 13 SIINTSQDFEKRIEIKKEVINPDVQF-----LEAHRAELTVAMIDEDLVLOEY 64
DB 742 SIQAVBEKMSNTEQLKQD-ISDAKTRQAEAMQDIKIEKQMSFNNK-DSKLAELQAT 799
QY 65 KQOQKHVDGKHFADCPNFVKGVPELYVDNNRIKIRYLQCPCKIKYDEERF-EABLITSH 123
DB 800 LBSLK-----KKLAKNSNSVVALQKELQIS-----RLSEQVGSDDLAAEQYAEAE----- 846
QY 124 HMRDITLAKLDIYVNHRRDLVYMAAADICTAITNGEYVKGILYGPFGTGSFILGA 183
DB 847 ---NTLKQWMEIOSKREQARV-KDAHDIAQOLEDERAK-----LTGFDEELRE 893
QY 184 IANQLSKKVRSTIYLPERITRLKGGFKDGSFEKTL-HRVREANTILMLDDIGA----- 236
DB 894 LEAATSKSKSRIT-----EEGLEMKLGHOLEK--LQKQQAQAQTVAH 935
QY 237 -EEVTPWVRDE 246
DB 936 MEEHEHMIADDE 946

Search completed: December 15, 2003, 15:19:02
Job time : 31.5681 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 / Search time 17.652 Seconds
(Without alignments)
1705.235 Million cell updates/sec

Title: US-09-689-952-2

Perfect score: 1643

Sequence: 1 MGSGSIMKQKSIINTSD.....ERVKSLSTPYFLSGENFRNN 313

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 9616682 residues

Total number of hits satisfying chosen parameters: 263308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

1: PIR 76:*

2: p1r1:*

3: p1r2:*

4: p1r3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1607	97.8	306	2	A89952
2	1631	38.4	311	1	I08544
3	601.5	36.6	311	2	H84042
4	552.5	33.6	307	2	AH1269
5	547.5	33.3	307	2	AB1632
6	407.5	24.8	298	2	B98066
7	403.5	24.6	293	2	H86718
8	392.5	23.9	298	2	D95199
9	173	10.5	209	2	S77882
10	145	8.8	282	2	D97138
11	145	8.8	316	2	B82907
12	138.5	8.4	235	2	E70378
13	137.5	8.4	399	2	G70328
14	133.5	8.1	313	2	D69945
15	131.5	8.0	329	2	D97340
16	128.5	7.8	739	1	A70204
17	127	7.7	266	2	I40411
18	126	7.7	287	2	AD1588
19	126	7.7	287	2	AG1651
20	126	7.7	440	2	D81415
21	120.5	7.3	316	2	T10436
22	117.5	7.2	267	2	E84092
23	117.5	7.2	449	2	A83650
24	117.5	7.2	454	2	D84931
25	116	7.1	842	2	G90576
26	115.5	7.0	246	2	D84932
27	115	7.0	246	2	C69197
28	114	6.9	265	1	BVEICIT
29	112.5	6.8	733	2	H69411

30	112	6.8	291	2	A86681	hypothetical prote
31	112	6.8	520	2	H69125	hypothetical prote
32	111.5	6.8	475	2	C96503	protein P9316.7 (1
33	111	6.8	350	2	B97735	probable Atfase n2
34	111	6.8	1208	2	T39068	coiled coil protel
35	110	6.7	461	2	A90512	chromosomal replic
36	110	6.7	866	2	E72113	clp proteinase ATP
37	110	6.7	866	2	H86508	ATP-dependent Clp
38	110	6.7	872	2	H81556	hypothetical prote
39	109.5	6.7	294	2	C86802	probable ABC trans
40	109	6.6	630	2	AC1309	endopeptidase Ia (
41	109	6.6	791	2	D81310	endopeptidase Ia (
42	109	6.6	831	1	A71825	replication initia
43	108.5	6.6	454	1	IQ4YBA	cell division cycl
44	108.5	6.6	759	2	D84301	cell division cont
45	108	6.6	811	2	B69512	

ALIGNMENTS

RESULT 1
A89952
primosomal protein [imported] - Staphylococcus aureus (strain N315)
C/Species: Staphylococcus aureus
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #ext_change 22-Oct-2001
C/Accession: A89952
R/Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; I
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramoto, K.
A/Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A/Reference number: A89758; PMID:21311952; PMID:11416146
A/Accession: A89952
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-306 <KUR>
A/Cross-references: GB:BA000018; PID:g13701480; PIDN:BA42774.1; GSPDB:GN00149
A/Experimental source: strain N315
C/Genetics:
A/Gene: dnaI
C/Superfamily: 44K dnaI protein homolog

Query Match 97.8%; Score 1607; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 5.4e-104;
Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	8	MGQFSIIINTSDPEKRIEIKIKEYINDPDVQKFLFAHRAELTNMIDEDLVNLOEYKQD	67
DB	1	MGQFSIIINTSDPEKRIEIKIKEYINDPDVQKFLFAHRAELTNMIDEDLVNLOEYKQD	60
QY	68	QKHYDGHKPADCPNFKVGHVPELVYDNNRIKIRYLQCPCKIRYDEERFAELITSHMQR	127
DB	61	QKHYDGHKPADCPNFKVGHVPELVYDNNRIKIRYLQCPCKIRYDEERFAELITSHMQR	120
QY	128	DTLNKLDIYNNHDDRLDVMAAADICGATITNGSVQGLYXGPRGKSTILGAINQ	187
DB	121	DTLNKLDIYNNHDDRLDVMAAADICGATITNGSVQGLYXGPRGKSTILGAINQ	180
QY	188	LKSKKVRSTIYLPFIRTLKGFQDGFPEKKLHVREANIIMLDIGAEVTPWVRDEV	247
DB	181	LKSKKVRSTIYLPFIRTLKGFQDGFPEKKLHVREANIIMLDIGAEVTPWVRDEV	240
QY	248	IGPLLHYRVNHELPTFFSSNFDYSLEHLAMTRDGEKTKARIIERYKSLSTPYFLSG	307
DB	241	IGPLLHYRVNHELPTFFSSNFDYSLEHLAMTRDGEKTKARIIERYKSLSTPYFLSG	300
QY	308	ENFRNN 313	
DB	301	ENFRNN 306	

RESULT 2

10B544
primosome component (helicase loader) dnaI - Bacillus subtilis
N:Alternate names: dnaI protein homolog, 44K, hypothetical protein Y (dnaB 3' region)
C:Species: Bacillus subtilis
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text-change 19-Jan-2001
C:Accession: B24720, C26580, F69617
R:Ogasawara, N.; Moriya, S.; Mazza, P. G.; Yoshikawa, H.
A:Title: Nucleotide sequence and organization of dnaB gene and neighbouring genes on the
A:Reference number: A53650; MUID:87117549; PMID:3027671
A:Accession: B24720
A:Molecule type: DNA
A:Residues: 1-311 <GUN>
A:Cross-references: GB:X04963; NID:G39880; PIDN:CA28633.1; PID:G39881
R:Hoshino T., McKenzie, T., Schmidt, S., Tanaka, T., Sueoka, N.
Proc. Natl. Acad. Sci. U.S.A. 84, 653-657, 1987
A:Title: Nucleotide sequence of Bacillus subtilis dnaB: a gene essential for DNA replication
A:Reference number: A94709; MUID:87118226; PMID:3027697
A:Accession: C26580
A:Molecule type: DNA
A:Residues: 1-18, 'N', 20-23, 'T', 25-206 <HOS>
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Aevedo, V.; Bertero
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Chica
C.; Ehrlich, S.D.; Emmericon, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gallen
leach, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.;
Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinio,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogilawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A.; Authors: Schlecht, S.; Schroeder, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpe, P.; Toogni, A.; Tosato, V.; Uchiyama,
T.; Winere, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:934377
A:Accession: F69617
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-311 <GUN>
A:Cross-references: GB:Z29118; GB:AL009126; NID:G2635200; PIDN:CA814858.1; PID:G2635363
A:Experimental source: Strain 168
C:Genetics:
A:Gene: dnaI
C:Superfamily: 44K dnaI protein homolog
C:Keywords: ATP, nucleotide binding, P-loop
P168-175/Region: nucleotide-binding motif A (P-loop)

[illegible]

```

RESULT 3
HB4042
primosome component (helicase loader) dnaI [imported] - Bacillus halodurans (strain C-12:
C1:Species: Bacillus halodurans
C1:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C1:Accession: HB4042
R1:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A1>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and s
A1:Reference number: AB3650; MUID:20512582; PMID:1105132
A1:Accession: HB4042
A1>Status: preliminary
A1:Molecule type: DNA
A1:Residues: 1-311 (Sto>
A1:Cross-references: GB:AB001517; GB:BA000004; NID:g10175500; PIDN:BAB06863.1; GSPDB:GN001
A1:Experimental source: strain C-125
C1:Genetics:
A1:Gene: dnaI
C1:superfamily: 44k dnaA protein homolog

```

```

Query Match 311, 36.6%, Score 601.5, DB 2, Length 311,
Beat Local Similarity 39.2%, Pred. No. 2e-34,
Matches 125, Conservative 69, Mismatches 104, Indels 21, Gaps 8,

Qy 5 QSIKQKFSIINTSQDFEKRLEKKEVINDDYKQFLPFAHRAELTNMIDEDLNLQEX 64
Db 2 ESISLSLKHMTG-NRSFEPQFTQLKEAFVRSFHVQLFLEEH-PSLSPIRLBGLSTLYEY 59
Qy 65 KDQQKHNDGH-KRADCPNFKGHVPELYVNNRKKIRKLOQPCCKIKYDEEFPEALLTS 122
Db 60 QKEQSHC-AHCPGLQKCPENLKKGYQPTLYVERDSLELSPPCLKEEREKKRSLIRS 118
Qy 123 HHMQDITNALTKDI-----YNNHRDLDPVMAADDTCTAITNGEQVKLYLPGPFT 175
Db 119 LYIKREILKAFPDVEDSESPGSIASHR-ALPFLALAKP-----GEDMGKLYLYKFGV 170
Qy 176 GKSPFLGAIANQLKSKKVRSTIYLPFPIRTLKGGFKDQSPFKLHRVREANILMDDIG 235
Db 171 GKTLMGAIANELDRGIDSTIIVVPDPFRFLKQSIGSTFGQKDLFVNAQVLIIPDDIG 230
Qy 236 AEEVTPVVRDEVIGPPLHYRMVHELPTFPSSNFDYSLEHHLAMT-RDGEKTKAARIIE 294
Db 231 AETWTSVRDDVLGVIILQYRIMEKLPPLFTSNYDDELEHILAINDKSGTELLKAKRVME 290
Qy 295 RVKSLSTPYFLSGENFRNN 313
Db 291 RIRHYTVSVVMYQGNRRHH 309

RESULT 4
AH1269
primosome component (helicase loader) DnaI [imported] - Listeria monocytogenes. (strain EC
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C/Accession: AH1269
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
. ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Feihl, H.;
Science 294, 849-852, 2001
D./ Jones, L.M.; Karet, U.
A./Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurupkat, G.; Medueno, E.; Maitournam, A.; Mat
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A./Title: Comparative genomics of Listeria species.
A./Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AH1269
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-307 <GLA>
A/Cross-References: GB:NC_003210; PIDN:CAC9638.1; PID:G16410989; GSPDB:GN00177
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: dnaI
C/Superfamily: 44k dnaA protein homolog

```

Query Match 33.6%; Score 552.5; DB 2; Length 307;
Best Local Similarity 37.8%; Pred. No. 4,7e-11;
Matches 113; Conservative 69; Mismatches 104; Indels 13; Gaps 4;

19 ODEKRIEIKKEVINDPDKOFLAHRALTNAMIDEDLNLOEYDQOQKHGKPAD 78
15 RDEKEVQGLKQOVLHOPIDPFKEKEBTEVLQVNVQNSLNYEFTQHKTEDE--- 71
79 CDFVKGAVELVYDNNRIKIRYLQCPCKIKYDERREALITSHMQRTLNAKLDIY 138
72 -ETLMGYPKVLNVEFIVTYVPTKEKIEBKRRRAVERIRISLVPKQVADNADFY 130
139 MNHRDLVMAADICTATITN-----GEQVKGILYVPGPTGKSPFLGAIANOLSKKY 193
131 TDESR---QALVAYOPLANTPPKSGEKVKGILFHSGFTGKSYLGLAKELAKGI 187
194 RSTIYLPFIRITLKGFGKDSFEKTLHVRVREANIIMLDIGAEVTPWVRDEVIGPLTH 253
188 STLVLPLPEFMRVKOSISDNTGKIQFAKETEVLMLDIGAESMTAMTRDEVILGIIQ 247
254 YKRVHLEPFPSSNFDYSELHHLAMTRDG-EKTKAARIIEVKSLSTPYPLSGENFR 311
248 FRMOEELPFPSSNFDYSELHHLAMTRDG-EKTKAARIIEVKSLSTPYPLSGENFR 306

RESULT 5

AB1632
primosome component (helicase loader) DnaI [imported] - *Listeria innocua* (strain Cl1p112)
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AB1632
R:Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maitounam, A.; Mak, C.; Schlueter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Weiland, A.; Title: Comparative genomics of *Listeria species*
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1632
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-307 <GAA>
A:Cross-references: GB:AL592022; PIDN:GAC96826.1; PID:G16414082; GSPDB:GN00178
A:Experimental source: strain Cl1p11262
C:Genetics:
A:Gene: dnaI
C:Superfamily: 44k dnaa protein homolog

Query Match 33.3%; Score 547.5; DB 2; Length 307;
Best Local Similarity 37.8%; Pred. No. 1e-30;
Matches 112; Conservative 68; Mismatches 109; Indels 7; Gaps 3;
19 ODEKRIEIKKEVINDPDKOFLAHRALTNAMIDEDLNLOEYDQOQKHGKPAD 78
15 RDEKEVQGLKQOVLHOPIDPFKEKEBTEVLQVNVQNSLNYEFTQHKTEDE--- 71
79 CDFVKGAVELVYDNNRIKIRYLQCPCKIKYDERREALITSHMQRTLNAKLDIY 138
72 -ETLMGYPKVLNVEFIVTYVPTKEKIEBKRRRAVERIRISLVPKQVADNADFY 130
139 MNHRDLVMAADICTATITN-NGEYVKGILYVPGPTGKSPFLGAIANOLSKKYRST 196
131 TDESRDLVAYOPLANTPPKSGEKVKGILFHSGFTGKSYLGLAKELAKGI 190
197 IYLPFIRITLKGFGKDSFEKTLHVRVREANIIMLDIGAEVTPWVRDEVIGPLTHYR 256
191 LVTLPEFMRVKOSISDNTGKIQFAKETEVLMLDIGAESMTAMTRDEVILGIIQFRM 250
257 VHELPFPSSNFDYSELHHLAMTRDG-EKTKAARIIEVKSLSTPYPLSGENFR 311
251 QEELPFPSSNFDYSELHHLAMTRDG-EKTKAARIIEVKSLSTPYPLSGENFR 306

RESULT 6

B98066
primosome component (helicase loader) [imported] - *Streptococcus pneumoniae* (strain R6)
C:Species: *Streptococcus pneumoniae*
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: B98066
R:Hoshino, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Buggett, S.; Dehoff, B.S.; Esy, P.; Jellison, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAnren, S.; Meyer, P.; Sun, P.M.; Winkler, M.B.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.; A:Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: B98066
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-298 <KUR>
A:Cross-references: GB:AB007317; PIDN:AAL00359.1; PID:G15459220; GSPDB:GN00174.
C:Genetics:
A:Gene: dnaI
C:Superfamily: 44k dnaa protein homolog

Query Match 24.8%; Score 407.5; DB 2; Length 298;
Best Local Similarity 32.0%; Pred. No. 4.8e-21;
Matches 94; Conservative 64; Mismatches 121; Indels 15; Gaps 7;

21 FEKRIEIKKEVINDPDKOFLAHRALTNAMIDEDLNLOEY-KDQKHGKPADC 79
17 YDVLQKIMK---DDVAFAVQ--QESINODELANSISKFNQYITERKFLG---DT 66
80 PNFVKGAVELVYDNNRIKIRYLQCPCKIKYDERREALITSHMQRTLNAKLDIY 139
67 DYIAKIKYKPLVNMHGVADVSTEEPELJAERKALIKRLNINPSSLKNVSLDYVR 126
140 MNHRDLVMAADICTATITNNGEYVKGILYVPGPTGKSPFLGAIANOLSKK-VRSII 198
127 DDVQRLTVLAKRIEFPNDPN--NLKGLVIXGDFGKSPFMAALAHLSKRGVSSTLL 184
199 YLEPFIITLKGFGKDSFEKTLHVRVREANIIMLDIGAEVTPWVRDEVIGPLTHYR 258
185 HXPSFVIDYKNAISDGNVTVLDEITLSDIIGAGQSTWVADDEILQVILQYRMGE 244
259 ELPTFPSSNFDYSELHHLAMTRDG-EKTKAARIIEVKSLSTPYPLSGENFR 311
245 NLPTFPSSNFDYSELHHLAMTRDG-EKTKAARIIEVKSLSTPYPLSGENFR 298

RESULT 7

H86718
primosomal protein DnaI [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL1403)
C:Species: *Lactococcus lactis* subsp. *lactis*
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: H86718
R:Belcetin, A.; Winkler, P.; Mager, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp
A:Reference number: A86625; MUID:2135186; PMID:11337471
A:Accession: H86718
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <STO>
A:Cross-references: GB:AB005176; PID:G12723668; PIDN:AAK04850.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: dnaI
C:Superfamily: 44k dnaa protein homolog

Query Match 24.6%; Score 403.5; DB 2; Length 293;
Best Local Similarity 32.3%; Pred. No. 9e-21;
Matches 101; Conservative 61; Mismatches 124; Indels 27; Gaps 8;

QY 8 MKQKSIINTSODEPEKLEIKKEVINDDPVKQFLAHRALTNAMIDEDLNTLOEX-KD 66
 DB 1 MESIGDILGKRDQIRENFEKLVAVLNKADVAQATIAQH--QMTDEORISKEYEYRE 58
 QY 67 QOKKYDHKFPADCFNFKGVHPELVYDNNRIKIRYLOCPCKIKYDEERFEALITSHH-- 124
 DB 59 HEKEKEKEKRA-----ADGVEPVLIMNHGVADSVS---QTNELAAQOQAQNLRRNNII 110
 QY 125 -MQRDTLNAKDKDLYMHRDLVYMAADDCTAITNGEYKGLYTGPFOTGSPFLGA 183
 DB 111 GLPFDLKVOTLADLADLVORIKRYQALVDF--ITNPPKKGLYLGDFGVGKSPMLAA 167
 QY 184 IANOLKSKVASTIYLPFIRTLKGGFKDGSPEKK--LHVRBANIMLDIGAEVIT 240
 DB 168 MANELAKKGISITLILHPTF-----SDLPFNAMKVMWEIRASQVLVLDIGAEQNN 220
 QY 241 PWRVDEIVGLLHYRMVHELPTFPSSNFDVSELEHMLAMTDEGEKTKAARIIVKSL 300
 DB 221 AMVRSILQVTLQHRMOENLPTFTSNLRMBELQHLAETGRADEIMPAKVMERVKYLA 280
 QY 301 TPYFLSGENFRNN 313
 DB 281 EEMRLBGTNRHD 293

RESULT 8

D95199
 C/Species: Streptococcus pneumoniae
 C/Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
 C/Accession: D95199
 R/Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidon, J.D.; Umeyam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radume, D.; Holtzaple, nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A/Authors: Lotfus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A./Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A/Reference number: A95000; MUID:21357209; PMID:11463916
 A/Accession: D95199
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-298 <KUR>
 A/Cross-references: GB:AE005672; PIDN:AAK5789.1; PID:gl4973206; GSPDB:GN00164; TIGR:SP4
 A/Experimental source: strain TIGR4
 C/Genetics:
 A/Gene: SP1711
 C/Superfamily: 44k dnaA protein homolog

Query Match 23.9%, Score 392.5; DB 2; Length 298;
 Best Local Similarity 32.3%, Pred. No. 5.3e-20;
 Matches 96; Conservative 64; Mismatches 116; Indels 21; Gaps 8;
 QY 21 FEKRIKIKKEVINDDPVKQFLAHRALTNAMIDEDLNTLOEX-KDQOKKYDHKFPAD 79
 DB 17 YQDLVQKIMK---DPVAAFIQ--QSLTPKEINRSISKNOYITERDKFLRG----DT 66
 QY 80 PNFYKGVHPELVYDNNRIKIRYLOCPCKIKYDEERFEALITSHHQRDTLNAKDKDLY 139
 DB 67 DYAKGVKPIVKNHGVADSVSEETPELIAEKEMAKNNRKLNTLPASLKSLAQQD 126
 QY 140 NHRDLVY---AMAAADICTAITNGEYKGLYTGPFOTGSPFLGAIANOLKSKV--VRS 195
 DB 127 DDLGRLLVPEKTLAFVQYPAIR-----KGLYLGDFGVGKSPVAAALADHLSKRGVSS 181
 QY 196 TIIVLPFIRTLKGGFKDGSPEKKLHVRBANIMLDIGAEVITPWRVDEIVGLLHYR 255
 DB 182 TLHAYPFAVLDVKNALSDGVNKTIVDEIKLSEVLILDDIGAEQSTVWVRDEIILQVILQYR 241
 QY 256 MVHMLPFPSSNFDVSELEHMLAMTDEGEKTKAARIIVKSLTPYFLSGENFR 311
 DB 242 MQENLPTFTSNFEDLEKHFACVKGNDETWEARRMERIRYLAETRLLEGVNR 298

RESULT 9

577882
 dnaA protein homolog - Mycoplasma capricolum
 N/Alternate names: protein MC072
 C/Species: Mycoplasma capricolum
 C/Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 07-Dec-1999
 C/Accession: S77882
 R/Bork, P.; Ouzounis, C.; Casati, G.; Schneider, R.; Sander, C.; Dolan, M.; Gilbert, W.; Mol. Microbiol. 16, 955-967, 1995
 A/Title: Exploring the Mycoplasma capricolum genome: a minimal cell reveals its physiology
 A/Reference number: S77739; MUID:96059641; PMID:7476192
 A/Accession: S77882
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-209 <BOR>
 A/Cross-references: EMBL:233058; NID:9514450; PID:CAA83732.1; PID:9530419
 A/Experimental source: ATCC 27343
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
 C/Genetics:
 A/Genetic code: SGC3

Query Match 10.5%, Score 173; DB 2; Length 209;
 Best Local Similarity 31.9%, Pred. No. 5.1e-05;
 Matches 51; Conservative 31; Mismatches 58; Indels 20; Gaps 8;

QY 108 IKYDERFEALITSHHQRDTLNAKDKDLYMHRDLVYMAADDCTAITNGEYKGL 167
 DB 57 LTYD-EYFKNEL---NNSFSLDQKRD---RKTL-----LSNLTKEVQKNSK-KGV 102
 QY 168 YLYGPFOTGSPFLGAIANOLKSKVASTIY--LPFIRTLKGGFKDGSPE--KKLHVR 223
 DB 103 YLYGHSIGTGYTFKVLANTLASK--NKTYIFSLRLIDKLKSEFSSSEINSLMKKI 160
 QY 224 REANIMLDIGAEVITPWRVDEIVGLLHYRMVHELPTF 263
 DB 161 KTVDFLFDLDIGENSLMARDDFLPELVNRYMENQATF 200

RESULT 10

D97138
 C/Species: Clostridium acetobutylicum
 C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
 C/Accession: D97138
 R/Nolling, U.; Breton, G.; Omeichenko, M.V.; Markova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
 A/Reference number: A96900; MUID:21359325; PMID:21359325
 A/Accession: D97138
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-282 <KUR>
 A/Cross-references: GB:AE001437; PIDN:AAK79895.1; PID:gl5024913; GSPDB:GN00168
 A/Experimental source: Clostridium acetobutylicum ATCC824
 C/Genetics:
 A/Gene: CAC1933

Query Match 8.8%, Score 145; DB 2; Length 282;
 Best Local Similarity 21.8%, Pred. No. 0.0065;
 Matches 51; Conservative 52; Mismatches 103; Indels 28; Gaps 7;
 QY 95 NRIRKIRLOCPCKIKYDEERFEA-----ELTSHHQRDTLNAKDKDLYMHRDL 145
 DB 44 DRIYKTMCEBSKAMHEKRQHEMLKKQNLNELITSLNDEKFRSNFPMWPNKSGRS 103
 QY 146 DVMAA---ADDI--CTAITNGEYKGLYTGPFOTGSPFLGAIANOLKSKVASTIYL 200
 DB 104 MYNLGRRYADNFECK-----QGLGLITGSAGNKGTIYASAIANELLQYIPVVCVSI 158
 QY 201 PEFIRTLKGGFKDGSPEKKLHVR--EANIMLDIGAEVITPWRVDEIVGLLHYRMV 257
 DB 159 NGLLSRIQKTYNSWGKEASDIIIRSFINDLLIIDLGTGTEKSSBSR-SMIVNIVDSRYR 217

QY 258 HELPTFSSNPDSYSELEHHLAMTRDGEKTKAARIIRVKSLSPPYLSGENFR 311
 DB 218 SKPLIITSLTLEINPSPKRGVLADQYHERIES-----RIEMCTPVENTSKSIR 266

RESULT 11

B82907
 conserved hypothetical ATP/GTP-binding protein U0307 [imported] - Ureaplasma urealyticum
 C:Species: Ureaplasma urealyticum
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: B82907
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.
 submitted to GenBank, February 2000
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit
 A:Reference number: A82870
 A:Accession: B82907
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-316 <GAA>
 A:Cross-references: GB:AE002128; GB:AF222894; NID:g6899279; PIDN:AAF30716.1; GSPDB:GN001
 A:Experimental source: serovar 3; biovar 1
 C:Genetics:
 A:Gene: U0307
 A:Genetic code: SGC3

Query Match 8.8%; Score 145; DB 2; Length 316;
 Best Local Similarity 21.3%; Pred. No. 0.0075;
 Matches 69; Conservative 57; Mismatches 130; Indels 68; Gaps 13;

QY 16 NTSODEKRIEIK-KEVINDPVK-----QFLRAHRAELTNAMIDSD-LN 59
 DB 3 NNOVDDELLENVYOKALEIPKAIINLTSEFPNLYFMDIVAYTSYLNANLNQDSY 62
 QY 60 VLOEYDQOKHYDGHKPADCPNFVKGHVPELYVNNRIKIRYLQCPCK-----IKYDE 112
 DB 63 IYOLNENEVNHL-----CLVVVK--KENKVD--KIKKNYILNITKNMDYNISLADDS 110
 QY 113 ERP--EAEILTSNM--QRPDLNAKLDIYMRDRDLVMAADICTAITNGEYKGLY 168
 DB 111 QIFSKKTEILDNLALERNKLINFLEARKKNKNTL-----IKNNNQCPKSAF 159
 QY 169 LVGPPFGSGFIIIGALANOLSKKVRSTIYLPEFIRTLKGGK-----DGSP 216
 DB 160 IYGDPEVGSIIITQATYNTI-SLKTKNLKIAVI-----TLMDLFKNVIQFPNYQDSDLYI 213
 QY 217 EKKLHVRREANITLMDDIGAEVTPWVRDEVIGPLHYRMVHSLPTFFSSNPDSYSELEH 276
 DB 214 NELVNLNSVDVVIDDFSSVNLNWSISITLMPITLNNRLKSTRQITFISNFSIEDLINS 273
 QY 277 LAMTRDGEKTKAARIIRVKSLS 300
 DB 274 TKNTNIEBOKTDLRFNRLEYLT 297

RESULT 12

E70378
 DNA replication protein DnaC - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C:Accession: E70378
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; NID:98196666; PMID:9537320
 A:Accession: E70378
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-235 <AGF>
 A:Cross-references: GB:AE000713; NID:g2983424; PIDN:AAC07013.1; PID:g2983431; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:

A:Gene: dnaC

Query Match 8.4%; Score 138.5; DB 2; Length 235;
 Best Local Similarity 23.8%; Pred. No. 0.014;
 Matches 59; Conservative 41; Mismatches 93; Indels 55; Gaps 9;

QY 82 FVKGHVPELYVNNRIKIRYLOCPCKIKYDEERFAELTSHMOBRLNAKIDYMNH 141
 DB 15 FVK-----TEDNKRLL-----CECFKRRDVREL-----NIPRYNNANI-DYHPK 56
 QY 142 RRLDVMAADICTAITNGEQVGLYLYPGFTGKSLIGALANOLSKK-VRSTIYL 200
 DB 57 NNSQNALTLIRVFVNFNPBECKGLTFVSGPGVKTHLAVATLKAIVEKGIKGYFDT 116
 QY 201 PEIRTLKGGFKDGSFEKTLHVRANITLMDDIGAEVTPWVRDEVIGPLHYRMVH 260
 DB 117 KDLIFRLKHMDEGCKTKFRTVNSPVLVDLGSERISDWQR-ELISYIIRYNNLK 175
 QY 261 PFPSSNPDSYSELEHHLAMTRDGEKTK-----AARIIRV----- 296
 DB 176 STIITNTY-----SLQREBSVRISADLASHGENVYSKTYENNELVYKSGDLR 226
 QY 297 --KSLSTP 302
 DB 227 KSKKSTP 234

RESULT 13

G70328
 chromosome replication initiator protein DnaA - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Sep-2000
 C:Accession: G70328
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
 V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; NID:98196666; PMID:9537320
 A:Accession: G70328
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-399 <AGF>
 A:Cross-references: GB:AE000683; NID:g2982996; PIDN:AAC06612.1; PID:g2983000; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: dnaA
 A:Superfamily: replication initiation protein dnaA

Query Match 8.4%; Score 137.5; DB 2; Length 399;
 Best Local Similarity 20.2%; Pred. No. 0.033;
 Matches 48; Conservative 52; Mismatches 69; Indels 69; Gaps 8;

QY 6 SIMKQFSIIINTSODEKRIEIKKE-----VINDPVQFLEAARAEITNAMIDEDLN 59
 DB 5 ALIKIESVDSYARELAKFE-IKQEKGFLEFLAPGEDYREMET-----IVNFFLEEV 59
 QY 60 VLOEYDQOKHYDGHKPADCPNFVKGHVPELYVNNRIKIRYLQCPCKIKYDEERFAEL 119
 DB 60 KLIIEVEK-----EKKKVEI 75
 QY 120 ITSHMQPRLTNAK--IKDIYNNHRLDLVMAADICTAITN-GRQVGLVLYGSPGTG 176
 DB 76 -----KQFLNPKTLENFVIGEGRL-----AYEVKALENLGLVNPFIYSGVGTG 124
 QY 177 KSFIIIGALINOLSKKVRSTIYLPEFIRTLKGGFKDGSFEKTLHVRANITLMDI 234
 DB 125 KTHLLQAAENAKKGGYRITYSSADDFQAAMVHKLKGTINFRMYNKSVDLLDDV 182

RESULT 14

D69945
 phage-related protein homolog yqam - Bacillus subtilis
 C:Species: Bacillus subtilis

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 9.90234 Seconds
(without alignments)
1486.453 Million cell updates/sec

Title: US-09-689-952-2

Perfect score: 1643

Sequence: 1 MCGGSGIMKQKFSINTSOD.....ERKSLSTPYFLSGENFRNN 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt 41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	631	38.4	311	1	DNAI_BACSU
2	137.5	8.4	399	1	O66569 bacillus su
3	133.5	8.1	313	1	YOAM_BACSU
4	128.5	7.8	457	1	DNAI_BACSU
5	127	7.7	266	1	XKOC_BACSU
6	126	7.7	440	1	DNAI_CAMDE
7	122	7.4	246	1	DNAI_BUCAP
8	117.5	7.2	449	1	DNAI_BACHD
9	117.5	7.2	454	1	DNAI_BUCAL
10	115.5	7.0	246	1	PSMR_BUCAL
11	115	7.0	410	1	PSMR_METTR
12	114	6.9	265	1	ISTB_PSEAE
13	112	6.8	445	1	DNAI_ACHLA
14	111.5	6.8	436	1	PSMR_METKA
15	111	6.8	1208	1	PCPI_SCHPO
16	109.5	6.7	229	1	ISTB_BURCE
17	109	6.6	791	1	LON_CAMDE
18	109	6.6	851	1	LON_HELPJ
19	108.5	6.6	434	1	DNAI_HELAP
20	107.5	6.5	263	1	ISTB_BACFR
21	107.5	6.5	1940	1	MYH3_CHICK
22	107	6.5	451	1	DNAI_PASMT
23	106.5	6.5	251	1	ISTB_BACST
24	105.5	6.4	231	1	ISTB_PSEFL
25	105.5	6.4	432	1	SKOI_SCHPO
26	105	6.4	248	1	YDAV_ECOLI
27	105	6.4	1938	1	MYH6_RAT
28	103.5	6.3	263	1	YAHM_RHISN
29	103	6.3	450	1	DNAI_MYCCA
30	102.5	6.2	466	1	DNAI_PSOIM
31	102.5	6.2	1939	1	MYH1_HUMAN
32	102	6.2	462	1	DNAI_YERPE
33	102	6.2	464	1	DNAI_SERMA

34	102	6.2	584	1	TRP4_YEAST
35	101	6.1	398	1	PSMR_ARCFU
36	101	6.1	402	1	PSR8_MANSE
37	101	6.1	810	1	YLZ5_CAEEL
38	101	6.1	1938	1	MYH6_MOUSE
39	100.5	6.1	777	1	LON_BUCAL
40	100.5	6.1	794	1	SC18_CANAL
41	100.5	6.1	1938	1	MYH3_CHICK
42	100	6.1	307	1	RYUB_MYCCE
43	100	6.1	405	1	PSR8_DROME
44	100	6.1	643	1	DNAI_STRRE
45	100	6.1	656	1	DNAI_STRCO

ALIGNMENTS

RESULT 1	ID	DNAI_BACSU	STANDARD	PRT	311 AA.
AC	P06567				
DT	01-JAN-1988	(Rel. 06, Created)			
DT	01-JAN-1988	(Rel. 06, Last sequence update)			
DT	28-FEB-2003	(Rel. 41, Last annotation update)			
DE	Primosomal protein dnaI.				
GN	DNAI.				
OS	Bacillus subtilis.				
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
OX	NCBI_TaxID=1423;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=87117549; PubMed=3027671;				
RA	Ogasawara N., Moriya S., Mazza P.G., Yoshikawa H.;				
RT	"Nucleotide sequence and organization of dnaB gene and neighbouring				
RL	genes on the Bacillus subtilis chromosome.";				
RN	Nucleic Acids Res. 14:9989-9999(1986).				
RP	[2]				
RX	SEQUENCE FROM N.A.				
RA	MEDLINE=97124191; PubMed=8969504;				
RT	Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,				
RL	Sanders J., Emerson P.T., Harwood C.R.;				
RT	"The dnaB-pha (256 degrees-240 degrees) region of the Bacillus				
RL	subtilis chromosome containing genes responsible for stress				
RT	responses, the utilization of plant cell walls and primary				
RL	metabolism.";				
RN	Microbiology 142:3067-3078(1996).				
RP	[3]				
RX	SEQUENCE FROM N.A.				
RA	STRAIN=168;				
RT	MEDLINE=98048467; PubMed=9387221;				
RL	Lapudis A., Galleron N., Sorokin A., Ehrlich S.D.;				
RT	"Sequencing and functional annotation of the Bacillus subtilis genes				
RL	in the 200 kb rmb-dnaB region.";				
RN	Microbiology 143:3431-3441(1997).				
RP	[4]				
RX	SEQUENCE FROM N.A.				
RA	STRAIN=168;				
RT	MEDLINE=98044033; PubMed=9384377;				
RL	Azevedo V., Bartero M.G., Bessieres P., Bolotin A., Borchert S.,				
RT	Kuntz F., Ogasawara N., Moser I., Albertini A.M., Allioni G.,				
RL	Borries R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,				
RT	Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,				
RL	Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,				
RT	Dentzot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,				
RL	Enutan K.D., Ertington J., Fabret C., Ferrari E., Foulger D.,				
RT	Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,				
RL	Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,				
RT	Gutsepri G., Guy B.U., Haga K., Haeck J., Harwood C.R., Henaut A.,				
RL	Hilbert H., Holsappel S., Hosono S., Hult M.F., Itaya M., Jones L.,				
RT	Joris B., Karamata D., Kasahara Y., Kletter-Blanchard M., Klein C.,				
RL	Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,				
RT	Kurita K., Lapudis A., Lardinois S., Lauber J., Lazarevic V.,				

RA Lee S.M., Levine A., Liu H., Maeda S., Mauei C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nishack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwolik S., Prescott A.M.,
RA Priescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha B., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorochin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terstra P., Tognoni A.,
RA Tosiato V., Uchiyama S., Vandebol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RT *subtilis*,"
RL Nature 390:249-256(1997).
RN (5)
RP SEQUENCE OF 1-206 FROM N.A.
RX MEDLINE=87118226; PubMed=3027697;
RA Hoshino T., McKenzie T., Schmidt S., Tanaka T., Sueoka N.,
RT "Nucleotide sequence of *Bacillus subtilis* dnaB: a gene essential for
RT DNA replication initiation and membrane attachment,"
RL Proc. Natl. Acad. Sci. U.S.A. 84:653-657(1987).
RN (6)
RP SIMILARITY TO DNAA.
RX MEDLINE=92195821; PubMed=1549481;
RA Koonin E.V.,
RT "Archaeobacterial virus SSV1 encodes a putative DnaA-like protein,"
RL Nucleic Acids Res. 20:1143-1143(1992).
RN (7)
RP IDENTIFICATION.
RX MEDLINE=95291463; PubMed=7773414;
RA Brund C., Ehrlich S.D.,
RT "The *Bacillus subtilis* dnaI gene is part of the dnaB operon,"
RL Microbiology 141:1199-1200(1995).
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: SOME, TO DNAA FROM VARIOUS BACTERIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04963; CAA28633.1; -
DR EMBL; 275208; CAA99605.1; -
DR EMBL; AF008220; AAC00359.1; -
DR EMBL; Z99118; CAB14858.1; -
DR EMBL; M5183; AAA22405.1; -
DR PIR; B24720; IQBS44.
DR Subtillet; BG10359; dnaI.
DR Interpro; IPR003593; AAA_Atpase.
DR SMART; SMO0382; AAA; 1.
KW Primosome; DNA replication; ATP-binding; Complete proteome.
FT NP_BIND 166 175
FT CONFLICT 19 19 K -> N (IN REF. 5).
FT CONFLICT 24 24 M -> T (IN REF. 5).
SQ SEQUENCE 311 AA; 36114 MW; A86FC9A86841264 CRC64;
Query Match 38.4%; Score 631; DB 1; Length 311;
Best Local Similarity 43.1%; Pred. No. 4.9e-36;
Matches 127; Conservative 64; Mismatches 102; Indels 2; Gaps 2;

QY 139 MNHRELVDVAMADICTAITNGGVKGLYLVPFGTGSFLIGAIYANOLKSKKASTII 198
DB 136 ISDPSTLAFQWHTPLDKSYNETGKKGKGLYLKGVGTPTFLAALANLAEKYSMTIV 195
QY 199 YLPEFRTLKGKFGKSGFEKRLHRYEANIIMDDIGAEVTPWVDEVIQPLHYRVA 258
DB 196 YLPEFRELKNSLDQDTLEKLNMTVTPTVLMDDIGAESMTSWVRDEVYGLQHRMSQ 255
QY 259 ELPTFFSSNPDYSELHHLAMTRDGE-EKTKAARIERKSLSTPYFLSGENRN 312
DB 256 QLPTFFSSNPDYSELHHLAMTRDGE-EKTKAARIERKSLSTPYFLSGENRN 310
RESULT 2
ID DNA_AQUAE STANDARD; PRT; 399 AA.
AC O66659;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chromosomal replication initiator protein dnaA.
GN DNA OR A0_322.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_Taxid=63363;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=VF5;
RC MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.B., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium *Aquifex*
RT *aeolicus*,"
RL Nature 392:353-358(1998).
RN (2)
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 76-399.
RX MEDLINE=2220049; PubMed=12234917;
RA Erzberger J.P., Pirruccello M.M., Berger J.M.,
RT "The structure of bacterial DnaA: implications for general mechanisms
RT underlying DNA replication initiation,"
RL EMBL J. 21:4763-4773(2002).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
CC (DNA BOX): 5'-TATTC(C/A)A/C(A/A)-3'. DNA BINDS TO ATP AND TO
CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE dnaA FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE00683; AAC06612.1; -
DR PIR; G70328; G70328.
DR PDB; 1L8Q; 25-SEP-02.
DR HAMAP; MF_00377; -; 1.
DR Interpro; IPR003593; AAA_Atpase.
DR Interpro; IPR001957; Bac_DnaA.
DR Pfam; PF00308; bac_dnaA; 1.
DR PRINTS; PR00051; DNAA.
DR SMART; SMO0382; AAA; 1.
DR TIGRFAMs; TIGR00362; DnaA; 1.
DR PROSITE; PS01008; DNAA; 1.
KW DNA replication; DNA-binding; ATP-binding; Complete proteome;
KW 3D-structure.
FT NP_BIND 119 126
FT CONFLICT 119 126 ATP (POTENTIAL).
SQ SEQUENCE 399 AA; 46840 MW; C641CB199AA79906 CRC64;


```

RA Paro V, Poll T.M., Portetelle D., Potwollik S., Prescott A.M.,
RA Presacco E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadie Y.,
RA Sato T., Scallan E., Schleich S., Schroeter R., Scoffone F.,
RA Setiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpeira P., Tognoni A.,
RA Tozato V., Uchiyama S., Vandenbol M., Vamler F., Vassarotti A.,
RA Viari A., Wandut R., Wedler E., Wedler H., Weitzneger T.,
RA Winters P., Wipit A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumestein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RL [4]
RP IDENTIFICATION.
RX MEDLINE=96084975; PubMed=7489895;
RA Medicine C., Moszer I., Viari A., Danchin A.;
RT "Analysis of a Bacillus subtilis genome fragment using a co-operative
RT computer system prototype.";
CC Gene 165:GC37-GC51(1995).
CC -I- SIMILARITY: TO B.SUBTILIS YOXC AND T.HYODYSENTERIAE HEMOLYSIN
CC TYA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-eb.ch/announce/
CC or send an email to license@isb-eb.ch).
CC -----
DR EMBL; D34216; BAA06927.1; -
DR EMBL; D84432; BAA12388.1; -
DR EMBL; Z99117; CAB14567.1; -
DR PIR; D69945; D69945.
DR Subtilist; BG11264; ygam.
DR InterPro; IPR003593; AAA_ATPase.
DR SMART; SM00382; AAA; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 313 AA; 36138 MW; 09208795E310D73A CRC64;
-----
Query Match 8.1%; Score 133.5; DB 1; Length 313;
Best Local Similarity 24.4%; Pred. No. 0.013;
Matches 68; Conservative 38; Mismatches 90; Indels 83; Gaps 15;
QY 60 VLQEVKQOQHNGHCKFA-----DCPNFVKGHVELVYDNNRIKIRYLQCPCKIKYD 111
Db 16 IMQRLRRSRARFLCTKQAVSEBKAFLDCP-----YCKRGIVYVRYNKHQTSWMLD 65
QY 112 EERFEALITSHNM--ORDTLNMLT-----KDIYNN-----HEDRLDVAAMAADIC 155
Db 66 E---QUDLMWPDMDVSSDDFLGKVCYCPDRAASEMKDITYSKQCEVCRRKRIARLMAASGIT 122
QY 156 ---TALING-----EQQYG-----LYYGPRTGKSPILG 182
Db 123 EEFKELFFGNFIITGKEDMIKDAYECAVEYYKDFOKIKGERONSIALLGQPGSKTHILLT 182
QY 183 AINQQL-KSKKYNSTIYLP--EIRITLKGSGFKGSGFEKLLHRRANIIIMLDI----- 234
Db 183 AINNNLLKKKSVH--CNYFFVBSMGDLKANF--DNEAKLKDARKKAVLEFIDDLFRPIN 238
QY 235 GAEVETPMVVRDEVIGPLLAHYRMVVELTFFSSNFDYSEL 273
Db 239 GQPRATDM-QVEQIQSVLNYRNYLNHKKRLILSSBELTIDEI 276
-----
RESULT 4
ID DNAA_BUCBP STANDARD; PRT; 457 AA.
AC P55567;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)

```

DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Chromosomal replication initiator protein dnaA.
 GN DNA OR BBP012
 OS Buchnera aphidicola (subsp. Baizongia pistacis).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Buchnera.
 RX NCBI_TaxID=135842;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2426901; PubMed=12522265;
 RA Van Ham R.C.H.U., Kamerbeek J., Palacios C., Raueell C., Abascal F.,
 RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.U.,
 RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
 RT "Reductive genome evolution in Buchnera aphidicola."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
 CC -1- FUNCTION: Plays an important role in the initiation and regulation
 CC of chromosomal replication. Binds to the origin of replication; it
 CC binds specifically double-stranded DNA at a 9 bp consensus (dnaA
 CC box): 5'-TATTC(C/A)A(C/A)A-3'. DnaA binds to ATP and to acidic
 CC phospholipids (by similarity).
 CC -1- SIMILARITY: Belongs to the dnaA family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AEO14016; AAO26756.1; -
 DR HAMAP; MF 00377; -; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS01008; DNaA; FALSE NEG.
 KW DNA replication; DNA-binding; ATP-binding; Complete proteome.
 FT NP BIND 162 169 ATP (POTENTIAL).
 SQ SEQUENCE 457 AA; 53365 MW; B2950813FADD4028 CRC64;
 Query Match 7.8%; Score 128.5; DB 1; Length 457;
 Best Local Similarity 22.3%; Pred. No. 0.046;
 Matches 67; Conservative 46; Mismatches 105; Indels 83; Gaps 12;
 QY 84 KGVPELVYDN-----NRKIRYLQ-----CP-----CKIYDEEPEAEI 119
 DB 34 KKNILILVAFNEFSFMKIKNYEINLKKLNKNINPTPTLMKICKPKIIOKKFNEI 93
 QY 120 ITSHNQDRTLAKLDIYNNHRDLVMAADICTAINTGEQVKG----- 166
 DB 94 T-----LKNILINSKLT--YNAVTKLSNIYSEINNTYFQNFQKQSNQLAFKIYKI 146
 QY 167 -----LVYGFPGFGKSPILGAINOLSKKVRSTIYL--PERIRLKGKFKD 213
 DB 147 AHNPGKNYFPLLYGSGSGKTHLHAVANNTIKYNTIKIYINSENFQMTSLKN 206
 QY 214 GSEFEKILHRYEANIIMLDDIGAEVTPWRDEV--IGLLHY-----RMVA 258
 DB 207 NTIEEPKRYKRVNTLLIDIOFPAYKKSQBELFTINMLLRNQCIIITSQPFQKH 266
 QY 259 ELPTFFSSNDYSELEHHLAMTDBGEKRYAARI---ERVKSLSTPY---PLSGENFRN 312
 DB 267 GIETRLKSRF-----ECGLTIRIDPDLNTRTKILKKSHIYDINLSKYVAFIANKLKS 321
 QY 313 N 313
 DB 322 N 322

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phase-like element PBX protein xkdc.
 GN xkdc
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 RX NCBI_TaxID=1423;
 RP [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=168 / SOL113;
 RA McDowell G.E., Wood H., Devine K.M., McConnell D.U.;
 RT "Genetic control of bacterial suicide: regulation of the induction of
 RT PBX in Bacillus subtilis."
 RL J. Bacteriol. 176:5820-5830(1994).
 CC [2]
 CC SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Krog S., O'Reilly M., Nolan N., Devine K.M.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 CC [3]
 CC SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berteiro M.G., Bessieres P., Bolechin A., Borchert S.,
 RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruchet C.V., Caldwell B., Capiano V., Carter N.M.,
 RA Choi S.K., Codani J.U., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Euteneier K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RA Fritz C., Glaser M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Gujara P., Goffeau A., Golightly E.J., Grandt G.,
 RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Bianhard M., Klein C.,
 RA Kobayashi Y., Koelter P., Koningsstein G., Krog S., Kumano M.,
 RA Kurita K., Lapdus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetlelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan B., Schleich S., Schroeder R., Scoffone F.,
 RA Sekiguchi J., Sekwaka A., Serror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,
 RA Takekichi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni K.,
 RA Tosato V., Uchiyama S., Vandembol M., Vannier F., Vassariotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weltzenegger T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zundstein E., Yoshikawa H., Zandhin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis."
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTIONAL ANTITERMINATOR.
 CC -1- SIMILARITY: STRONG, TO B. SUBTILIS Y0AM.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z34287; CAA84044.1; -
 DR EMBL; Z70177; CAA84054.1; -
 DR EMBL; Z99110; CAA83110.1; -
 DR PIR; I40411; I40411.
 DR Subtilist; BG10996; xkdc.
 DR InterPro; IPR003593; AAA_ATPase.
 DR SMART; SM00382; AAA; 1.
 KW ATP-binding; Complete proteome.

```

SQ  SEQUENCE 440 AA; 49686 MW; A3FAB90CCED06BD7 CRC64;
Query Match 7.7%; Score 126; DB 1; Length 440;
Best Local Similarity 21.3%; Pred. No. 0.067;
Matches 64; Conservative 51; Mismatches 93; Indels 92; Gaps 13;
QY 0PTEAHAEYTNMADIDLVLCGYKQOQKHGDKFADCPNFKGVPELYVN-NRI 97
DB 5 QILENKKKEISENEYEYLSNLKNEQSK-----ADLVNAEREL 46
QY 98 KIRYLQCPCKIKYDEERFEAEILITSHMORDTLNAKLKI-----TMNHRDLVA--- 148
DB 47 MAKRIQ-----TKYQKK-----IAHYEVQSGKKAIIINQASAKQSNSTKIDIAHIK 95
QY 149 -----MAADICTAITNGEYQKGLY-----LYGPEGTGKSPILGA 183
DB 96 AOSTIILNPSPEFESFVVGDSNKKYAVGACKAIHAHDKIGKLKYNPIFYVGPGLGKTHLLQA 155
QY 184 IANQKSKKXRTIIV--PEPIRTLKGGRDGSFEKKLHRYRANILMTDIGAEEVTP 241
DB 156 VGN--ASLEWKGKVIYATSENFINDFTSNLKNSLDKFKHEKRYKCDVLLIDYVFLKTD 213
QY 242 WVRDEVIGPLIHYRMVHELPTFFSSNDVSELEH--LMTNRGEEKTAARIIEKVS 298
DB 214 KIQE-----FF--FIENEIKNDDQIIMTSINPMLKG-ITERLKS 253

RESULT 7
DNAC_BUCAP
ID DNAC_BUCAP STANDARD; PRT; 246 AA.
AC Q8KA79;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA replication protein dnac.
GN DNAC OR BUGS022.
OS Buchnera aphidicola (subsp. Schizaphis graminum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxId=98794;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22084549; Pubmed=120899438;
RX Tamas I., Klassen L., Cambaeck B., Naeslund A.K., Eriksson A.-S.,
RA Wernegreen J.J., Sanderstrom J.P., Moran N.A., Andersson S.G.E.;
RL "50 million years of genomic stasis in endosymbiotic bacteria.";
Science 296:2376-2379(2002)
CC -1- FUNCTION: THIS PROTEIN IS REQUIRED FOR CHROMOSOMAL REPLICATION. IT
CC FORMS, IN CONCERT WITH DNAB PROTEIN AND OTHER PREPRIMING PROTEINS
CC DNAT, 'N', 'N', 'A PREPRIMING PROTEIN COMPLEX ON THE SPECIFIC SITE
CC OF THE TEMPLATE DNA RECOGNIZED BY PROTEIN 'N' (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE DNAC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements)
CC or send an email to license@ebi.ac.uk.
CC -----
CC EMBL; AEO14077; AAM67594.1;
CC DNA replication; Primosome; Complete proteome.
CC SITE 69 69 PROBABLY INVOLVED IN THE INTERACTION
CC WITH THE DNAB PROTEIN (BY SIMILARITY).
SQ SEQUENCE 246 AA; 28497 MW; BFE7E2A9BC2ADBBD5D CRC64;
Query Match 7.4%; Score 122; DB 1; Length 246;
Best Local Similarity 23.6%; Pred. No. 0.064;
Matches 48; Conservative 39; Mismatches 86; Indels 30; Gaps 7;
QY 111 DEERFEAEILITSHH---WQDTLNAAKLDIYNN-----HRDLVYMAAADICTA 157

```

DB 35 EGRSLSESSILRENMKQVILGRSGIRELYNMNCSPENYKIEHGCRKYLKAAKRYAE 94
 QY 158 ITNGEOVKGLYLVPGFTGKSPFLIGAIANOLKSKVSTIIYLPFIRTLKGGFKDS-- 215
 DB 95 P--NENIASIFSGRPGTGKHLASAGNLIHLGSKILVTYADLSNKKGFTSGTSNI 152
 QY 216 -FEKQHRVREANIMLDDIGAEVTPWVDEVIPLLHYRMHELTPFSSNPDYSELE 274
 DB 153 TEENLHNLSSVDLMDEIGMGTESRYEK-VIINQVDRSSSKSGTGLMSLNDHRGMK 211
 QY 275 HHLAMTRDGEKTKAAIIIRVK 297
 DB 212 NLL-----GR-----RVIDRR 223
 RESULT 8
 DNAA_BACHD STANDARD; PRT; 449 AA.
 ID DNAA_BACHD STANDARD; PRT; 449 AA.
 AC Q9RCZ2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chromosomal replication initiator protein dnaA.
 GN DNAA OR BH0001.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=99356711; PubMed=10427704;
 RA Takami H., Masui N., Nakasone K., Horikoshi K.;
 RA Fujii F., Hitama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
 OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
 IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
 (DNA BOX): 5'-TATATC(C/A)A(C/A)-3'. DNAA BINDS TO ATP AND TO
 ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the dnaA family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: AB013492; BAA82685.1; -
 DR EMBL: AP001507; BAB03720.1; -
 DR PIR: A83650; A83650.
 DR HAMAP: MF_003777; - 1.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001957; Bac_DnaA.
 DR Pfam: PF00308; bac_dnaa; 1.
 DR PRINTS: PR00051; DNAA.
 DR SMART: SM00382; AAA; 1.
 DR TIGRFAme: TIGR00362; Dnaa; 1.
 DR PROSITE: PS01008; DNAA; 1.
 DR KW DNAA replication; DNA-binding; ATP-binding; Complete proteome.
 FT NP_BIND 153 160

SQ SEQUENCE 449 AA; 51263 MW; D75C0BBF110A614 CRC64;
 Query Match 7.2%; Score 117.5; DB 1; Length 449;
 Best Local Similarity 20.8%; Pred. No. 0.28;
 Matches 67; Conservative 49; Mismatches 107; Indels 99; Gaps 16;
 QY 21 FEKRIKIKKEVINPDV-----KQFLAHRALTNAMIDEDLMTLQAEYKDOOKHY 71
 DB 24 FETMLKQTANSIEDSTIIITAVNEFARQWLBEHYDELLSETID-DLTGVRLV----- 75
 QY 72 DGHKPADCPNFKVGHVPELVYDNNRIKIRYLQCPCKIKIDEFEFAELITSHMQRDTLN 121
 DB 76 -----PKFV-----IPTSLD-----EPFVQELKKPKQPPAQN 105
 QY 132 AKLQDIYNNHRLDVA-----ADDICTATNGEOVKGLYLVPGFTGKSPFLIGA 183
 DB 106 GEMPNMMLNDKTFDFVIGSGNRFHAAASLAVAPAKAYNPLFYGGVGLKTHLMHA 165
 QY 184 IANQLSKKVRSTIIYLP-----EFIRTLKGGFKDSFEKQLHVRANIIMLDDI--- 234
 DB 166 IGHVMDHNPNAKVYVLSSEKFTNEFTNALRDN-KAVNFRNK---YRNVDLLIDDIQFL 221
 QY 235 -GAEEVTPWVDEVIQPL-LHY-----RMVHELPTF--FSSNPDYS----- 271
 DB 222 AGKEQ-----TOEEFFFTNALHEDNKOIYISSDPKKEIPTLEDRLSRFEWGLITDITP 277
 QY 272 -ELHHLAMTRDGEKTKAARI 292
 DB 278 PDLERITAIIR--KKAQENL 296

RESULT 9
 DNAA_BUCAI STANDARD; PRT; 454 AA.
 ID DNAA_BUCAI STANDARD; PRT; 454 AA.
 AC P57128;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chromosomal replication initiator protein dnaA.
 GN DNAA OR BU012.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tokyo 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. APS.";
 RL Nature 407:81-86(2000).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
 OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
 IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
 (DNA BOX): 5'-TATATC(C/A)A(C/A)-3'. DNAA BINDS TO ATP AND TO
 ACIDIC PHOSPHOLIPIDS.
 CC -1- SIMILARITY: Belongs to the dnaA family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: AP001118; BAB12740.1; -
 DR HAMAP: MF_003777; - 1.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001957; Bac_DnaA.
 DR Pfam: PF00308; bac_dnaa; 1.

DR PRINTS; PRO0051; DNAA.
 DR SMART; SM00382; AAA; 1.
 DR TIGR00362; DNAA; 1.
 DR PROSITE; PS01008; DNAA; 1.
 DR DNA replication; DNA-binding; ATP-binding; Complete proteome.
 KW NP BIND 159 166 ATP (POTENTIAL).
 SQ SEQUENCE 454 AA; 52931 MW; 1F5299FAC2751213 CRC64;

Query Match 7.2%; Score 117.5; DB 1; Length 454;
 Best Local Similarity 22.1%; Pred. No. 0.28;
 Matches 53; Conservative 46; Mismatches 86; Indels 55; Gaps 10;

QY 25 IRIKKEVINDPDKVQFLBAHRAELTNAMIDEDLVYQEKDQKKYDGHKFAQDCEPV 83
 DB 10 LRLQDELN-TEFSWIRSLKAKLN-----NILEY-----ANKKV 47
 QY 84 KGHVPELVYDNNRIKIRYQPC-----KIKYDEEFPEALITSHMQDITNAKIKDIY 138
 DB 48 LEWVKDKYL-NHLK-KILODYCGTNSPLIKFEIYQIYENKDKKNIENNKKNEKLIW 104
 QY 139 MN-----HNRDLVMAAADICTAINT-----GEQVGLYIYGPFG 176
 DB 105 SNIPKKNLSYRNINRKNYFQNFVEGSKNQLARSAFOAARNPGNSYNPLFYGTGIG 164
 QY 177 KSFILGAINOLSKKVRSTIYLPF--FIRTKGFGKDSFEKKLHVRVANIILMDI 234
 DB 165 KTHLMAIGNEIISYKIDIKIIPNSECFQDMVKALKNNATKFKLYRSDALIIDI 224

RESULT 10
 DNAC BUCAL STANDARD; PRT; 246 AA.

AC P57J34;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA replication protein dnac.
 GN DNAC OR BU021.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxId=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Tokyo 1998;
 RX MEDLINE=20445173; PubMed=109933077;
 RA Shigenobu S., Matanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. APS";
 RL Nature 407:81-86(2000).
 CC -1- FUNCTION: THIS PROTEIN IS REQUIRED FOR CHROMOSOMAL REPLICATION. IT
 FORMS, IN CONCERT WITH DNAB PROTEIN AND OTHER PREPRIMING PROTEINS
 DNAT, N, N', A PREPRIMING PROTEIN COMPLEX ON THE SPECIFIC SITE
 OF THE TEMPLATE DNA RECOGNIZED BY PROTEIN N' (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DNAC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF001118; BAB12748.1;
 CC DNA replication; Primosome; Complete proteome.
 FT SITE 69 69 WITH THE DNAB PROTEIN (BY SIMILARITY).
 SQ SEQUENCE 246 AA; 28444 MW; C2E56A3BCD1A68F CRC64;

Query Match 7.0%; Score 115.5; DB 1; Length 246;
 Best Local Similarity 22.5%; Pred. No. 0.19;

Matches 49; Conservative 40; Mismatches 88; Indels 41; Gaps 8;

QY 107 KIKYDEE-----RFEALITSHN-----MORDITNAKLIYNN-----HR 142
 DB 20 KRFPNDEDLAMNOEGGLSSSESTIREKKMKQVRIGRSIRRELYNMSFDPNKEYED 79
 QY 143 DRLDVAAMADICTAINTGEQVGLYIYGPFGTGSFILGAIANOLSKKVRSTIYLPF 202
 DB 80 GGRKVLKASKRYAEF--NENIASFISGKPGTGNHLSAIGNYILHKGSIILVTAD 137
 QY 203 FIRTGKFGKDS---FEKKLHVRVANIILMDIGAEVTVWVDEYIGPLHVRVH 259
 DB 138 LMSNNKMGFGSGNSNTEETELHLDLSSVDLMDIDEGMQTESRYEK-VIIINQIVDRSSSK 196
 QY 260 LPTFPSSNDPSELEHNLAMTRDGEKTKAARIIIRVK 297
 DB 197 RSTGMLSNIDHKMKSLL-----GR-----RVIDRMR 223

RESULT 11

PSNR_METH STANDARD; PRT; 410 AA.
 ID PSNR_METH
 AC O26824;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Proteasome-activating nucleotidase (Proteasome regulatory subunit).
 GN PAN OR MTH728.
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxId=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=93711463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Potter B., Qiu D.,
 RA Spadafora R., Vicare R., Wang Y., Wleczkowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Sater H., Patwell D., Prabhakar S.,
 RA McDougall S., Shiner G., Goyal A., Pletrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- FUNCTION: Required for the ATP- or GTP-dependent degradation of
 CC proteins, but not small peptides, by the 20S proteasome (By
 CC similarity).

CC -1- SUBUNIT: Homohexamer (Potential).
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF000852; AAB85233.1;
 CC PIR; G69197; G69197.
 CC HAMAP; MF_00553; -1.
 DR InterPro; IPR005937; 268_p45.
 DR InterPro; IPR003593; AAA_Arpase.
 DR InterPro; IPR003959; AAA_Arpase_cent.
 DR InterPro; IPR003960; AAA_sub.
 DR Pfam; PF00004; AAA_1.
 DR SMART; SM00382; AAA; 1.
 DR TIGR00362; DNAA; 1.
 DR PROSITE; PS00674; AAA; 1.
 KW Proteasome; ATP-binding; Complete proteome.

FT NP BIND 192 199 ATP (POTENTIAL).
 SQ SEQUENCE 410 AA; 4611 MM; 38261703F836AD4 CRC64;
 Query Match 7.0%; Score 115; DB 1; Length 410;
 Best Local Similarity 22.3%; Pred. No. 0.37;
 Matches 59; Conservative 50; Mismatches 101; Indels 54; Gaps 12;

16 NTSDQFKEKRIEKEV--INDPDKQ---PLEAHEAEITNMAIDDLNVL--QEVYDQ 68
 3 NNSQNVAKIEDLKCEYRLMKESKTKRNLMLKRLKLEQDKLIERRLDRVYSLR 62
 69 KHYDGHKPADCPNFKGVHPELVYNNRIKIRYQCP--CKIKYDEEFEEALITSHMOR 127
 63 GEIEREFP---TPPLVITVTEV--LDDHRAVAVKSTGPHFVINYSRFIDRKQLEGGARVAL 118
 128 DTLNAKDKDLYMHNRDLVDVMAAD---DICTAITNG--EYV----- 164
 119 NQOTFSDVDLPSPKDDVVTGMEVEKPDVSYEQIGLEQVREKVTLPKPELFE 178
 165 -----KGLVYGPFGTSGKSPILGAIANOLKS---KKVSTII--YLPEFIRTLKGKGF 212
 179 KIGIEPPKGVLTICPCTGTILAKAVAHETNAFIIKIVASEFVKYKIGGARLVKGVF-- 237
 213 DGSFEKLRHVRREANIIMDDIGA 236
 238 -----ELAKEKSPSIIIFIDEIDA 255

Db

RESULT 12
 ISTD_PSEAB STANDARD; PRT; 265 AA.
 AC P15026;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Insertion sequence IS21 putative ATP-binding protein.
 GN ISTD.
 OS Pseudomonas aeruginosa.
 OC Plasmid R68.45.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89218951; PubMed=2540414;
 RA Reilmann C., Moore R., Little S., Savioz A., Willetts N.S., Haas D.;
 RT "Genetic structure, function and regulation of the transposable
 element IS21.";
 RL Mol. Gen. Genet. 215:416-424(1989).
 RN [2]
 RP REVISION TO 283.
 RA Berger B.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE ISTD PROTEIN, ONE OF TWO PROTEINS EXPRESSED ONLY
 CC WHEN THERE IS A TANDEM REPEAT OF THE IS21 INSERTION SEQUENCE, IS
 CC NECESSARY FOR THE TRANSDUCTION OF PLASMIDS WITH THAT TANDEM
 CC REPEAT.
 CC -1- SIMILARITY: BELONGS TO THE IS21/IS1162 PUTATIVE ATP-BINDING
 CC PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X14793; CA33899.2; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR002611; ISTD_ATPbind.
 DR Pfam; PFO1695; ISTD; 1.
 DR SMART; SM00382; AAA; 1.

KW Plasmid; Transposable element; ATP-binding.
 FT NP BIND 104 111 ATP (POTENTIAL).
 SQ SEQUENCE 265 AA; 30528 MM; 904C7ADC5E3A12A9 CRC64;
 Query Match 6.9%; Score 114; DB 1; Length 265;
 Best Local Similarity 26.1%; Pred. No. 0.26;
 Matches 42; Conservative 36; Mismatches 53; Indels 30; Gaps 8;

162 EOVKGYLYGPGTSGKSPILGAIANOLKSKKRVSTIYLP--EFIRTLKGFGDGSPEKK 219
 95 ERSENVILGPPGVGKTHL--AIALGVKAVDAGRVLVFLPDLRLITLTKAKQENRLERQ 152
 220 LHRREANILIMDDIGAEEVTPWVRDEVIQPLAHYMH---ELPTFESSN--PDYSE 272
 153 LQQLSYARVILIDEIG---YLPWNRREA---SLFPLRLNRRYKASIIILSNKGADWGE 206
 273 L--EHLNLTROGEKTKARIIRKYSISTPFLSGENR 311
 207 MFGDHVL-----ATAIDRLHLHSTLTINIGESYR 236

Db

RESULT 13
 DNAA_ACHLA STANDARD; PRT; 445 AA.
 ID DNAA_ACHLA
 AC Q9KRD8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chromosomal replication initiator protein dnaA.
 GN DNAA.
 OS Acholeplasma laidlawii.
 OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
 OC Acholeplasmataceae; Acholeplasma.
 OX NCBI_TaxID=2148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Pg-8B;
 RX MEDLINE=20242229; PubMed=10779957;
 RA Taganov K.D., Gushchin A.E., Akopian T.A., Oparina N.Y.,
 RA Abramychcheva N.Y., Govorun V.M.;
 RT "Analysis of genes, coding for DNA gyrase from the mycoplasma
 RT Acholeplasma laidlawii Pg-8B.";
 RL Mol. Biol. (Mosk) 34:292-299(2000).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE INITIATION AND REGULATION
 CC OF CHROMOSOMAL REPLICATION. BINDS TO THE ORIGIN OF REPLICATION;
 CC IT BINDS SPECIFICALLY DOUBLE-STRANDED DNA AT A 9 BP CONSENSUS
 CC (DNAA BOX): 5'-TTATC(C/A)A(C/A)A-3'. DNAA BINDS TO ATP AND TO
 CC ACIDIC PHOSPHOLIPIDS (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the dnaA family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AP248639; AAF75986.1; -;
 DR HAMAP; MF_00377; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001957; Bac_DnaA.
 DR Pfam; PFO0308; bac_dnaA; 1.
 DR PRINTS; PR00051; DNAA.
 DR SMART; SM00382; AAA; 1.
 DR TIGRPFAM; TIGR00362; DnaA; 1.
 DR PROSITE; PS01008; DNAA; 1.
 KW DNAA replication; DNA-binding; ATP (POTENTIAL).
 FT NP BIND 144 151
 SQ SEQUENCE 445 AA; 51046 MM; A6045E53F8C9DB3 CRC64;
 Query Match 6.8%; Score 112; DB 1; Length 445;
 Best Local Similarity 22.6%; Pred. No. 0.67;

Matches 44; Conservative 36; Mismatches 85; Indels 30; Gaps 9;

QY 58 LNVLYQYKQOQKHYPDHPKADCPN-FYKGVHVPRLAYDN-NRIRIYLCQPCCKIKY-DEER 114
 Db 29 LPVTSFVKOD---NGLITWVANEFLKRNINLCTAKINELATKYSSSTVRLKFPVSOER 84
 QY 115 FEAELTSHMORDTLNKLKDIY-----MNRDRDLVMAAADICTAINTSEQVKG 166
 Db 85 VIEEPVADRRLTIDYQGNLNTSTYTPDSFVVGKSNMFAFMAMKADHPQAVAN-----P 139
 QY 167 LYIYGPFGTGSFLGAINOLSKKYRSTIYIY-----LPEIRTL-KGPFQGSFEKK 219
 Db 140 FYIFGVGVGKTHLMQALIGVYIINDVYKRLVYKADNFIEDVSLSRKNTKEEFNAK 199
 QY 220 LHRVREANITMLDDI 234
 Db 200 ---YKQIDVILVDDI 211

RESULT 14
 PSMR METKA STANDARD; PRT; 436 AA.

AC Q8TX03;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Proteasome-activating nucleotidase (Proteasome regulatory subunit).
 GN PAN OR MK0878.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyrii; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxId=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
 RA Shchepochina O.V., Shakhova V.V., Belova G.I., Atavind L.,
 RA Natile D.A., Rogozin I.B., Tatunov R.L., Wolf Y.I., Stetter K.O.,
 RA Malykh A.G., Koonin E.V., Kozlyavkin S.A.,
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 and monophyletic genome of archaeal methanogens.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 CC -1- FUNCTION: Required for the ATP- or CTP-dependent degradation of
 proteins, but not small peptides, by the 20S proteasome (by
 similarity).
 CC -1- SUBUNIT: Homohexamers (Potential).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; A5010377; AAM02091.1; -
 DR HAMAP; MF 00553; -; 1.
 DR InterPro; IPR005937; 26S_P45.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003959; AAA_ATPase_centre.
 DR InterPro; IPR003960; AAA_sub.
 DR Pfam; PF00004; AAA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRPFAM; TIGR01242; 26SP45; 1.
 DR PROSITE; PS00674; AAA; 1.
 DR Proteasome; ATP-binding; Complete proteome.
 KM NP_BIND 220 227 ATP (POTENTIAL).
 FT SEQUENCE 436 AA; 49767 MW; A515008616A02ED CRC64;
 Query Match 6.84; Score 111.5; DB 1; Length 436;

Best Local Similarity 21.5%; Pred. No. 0.71;
 Matches 59; Conservative 55; Mismatches 100; Indels 61; Gaps 13;

QY 5 QSIMKO-FKSIINTSODFEKRIIEKIGVYINDPVQKPIFAHR--AELTNMIDDLAV 60
 Db 27 BELIKYFRLERLEKRLRAHEKRLIEARRRRRLKEKLEMEDEKALREELRRREVM 86
 QY 61 LQEVYKQOQKHYPDHPKADCPN-FYKGVHVPRLAYDN-NRIRIYLCQPCCKIKY-DEER 114
 Db 87 EKIRSLQF-----MKRPLVIGTVEI-LDDGVYKSSGPFVSNVSTVDRL 138
 QY 118 ELTSHMORDTLNKLKDIY-----MNRDRDLVMAAADICTAINTSEQVKG 166
 Db 139 EPGAVNALNQSM-AVVDVLPSEKOSRVLAWEVDSPVSYDDIGLEQREIREV 196
 QY 165 -----KLYIYGPFGTGSFLGAINOLSKKYRSTIYIY-----LPEIRTL 207
 Db 197 KPLKEPELFKGVVPEPKGVLLYGPPTGKTLLAKVANHAADATFIR---LAAPELVORF 253
 QY 208 KGFQDGS-----PEKLRVREANITMLDDIGA 236
 Db 254 IG---EGARLVRELF--LAREKAPSLIFIDEIDA 293

RESULT 15
 PCPI SCHPO STANDARD; PRT; 1208 AA.

AC Q92351;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Spindle pole body protein pcpi.
 GN PCPI OR SPAC693.06C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OX NCBI_TaxId=4896;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, AND SUBCELLULAR LOCATION.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Baeham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Felkewell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Nibbelk D., Odeil C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grympey B.,
 RA Wellens I., Vansireels B., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck C., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,


```

RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: Spindle pole body component that binds calmodulin.
CC Overexpression of pcpl causes the formation of supernumerary SPB-
CC like structures and disrupts both mitotic spindle assembly and
CC chromosome segregation.
CC -1- SUBCELLULAR LOCATION: Spindle pole body.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, Z81317; CAB03608.1; -.
DR EMBL, AF348506; AAK31344.1; -.
DR PIR, T39068; T39068.
DR GeneDB Spombe; SPAC6G9.06c; -.
DR Colled_coll.
KM Colled_coll.
FT DOMAIN 151 375 COILED COIL (POTENTIAL).
FT DOMAIN 387 803 COILED COIL (POTENTIAL).
FT DOMAIN 874 1091 COILED COIL (POTENTIAL).
FT DOMAIN 1177 1204 COILED COIL (POTENTIAL).
SQ SEQUENCE 1208 AA; 140763 MW; 70264159ADDA42424 CRC64;

Query Match 6.8%; Score 111; DB 1; Length 1208;
Best Local Similarity 20.6%; Pred. No. 2.6;
Matches 72; Conservative 52; Mismatches 113; Indels 112; Gaps 16;

QY 4 GOSIMKOPKSIINTSQDFKRIEKIKKEVINDPVKQFLBAHRAELTNAMIDE----- 56
DB 693 GQKLASAFELMSNEKQALSKYSLKNELINAQNLDRREBELSELKCLFEERKTRSGS 752
QY 57 -----DIANL-QEYKD--QOKHYDGKFA--DCPNFVKGHVPELYVDNNRIKIRY- 101
DB 753 NNDIEKKKEINVLNSELADKLAAQIRHLESDKMELDKLVHHLNNGIEBANIEENAVKRLC 812
QY 102 LQPC-----KIKYDEERF---EALITSHHMQR 127
DB 813 LLMGCDYSSVSIQIYVSIQIEHFVNOQIQTRSLKQELRHDFVQPSGKKEQL--SRSFEK 870
QY 128 DTINAKLKDIYNNHRDLDAVMAADICTA---ITNGEQVKGILYGFPGTGSFTIGAI 184
DB 871 FGIETETKHIIAQRNR-NVSEKNDLENAAQKPFSSPDRKNGLYLPSHTSK----- 922
QY 185 ANQKSKKVRSTIIYLEPIRTLKGFKDGSPFEKQLHVRBANILMDDIGA----- 236
DB 923 -----IEYEKTIEDLKALQD-----ELKRNILM-DDISSYKQTTKL 961
QY 237 EYVTPWRDEVIGPLHYRWVHELPTFSSNPDY-----SELEHNL 277
DB 962 QEKIKWLERE-----RSILIDELSEYSRNOFNQNNLVQDKNELLEERL 1004

```

Search completed: December 15, 2003, 15:14:17
 Job time: 10.9023 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 15:11:40 ; Search time 36.1681 Seconds
(without alignments)
2233.384 Million cell updates/sec

Title: US-09-689-952-2

Perfect score: 1643
Sequence: 1 MGCGGQIMQFKSIINTSD.....BRVKSISTPYFLSGENFRNN 313

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1607	97.8	306	16	Q99TH8
2	1599	97.3	306	16	Q8NM67
3	1376	83.7	306	16	Q8CNV2
4	621.5	37.8	308	16	Q8CXB0
5	601.5	36.6	311	16	Q9K863
6	552.5	33.6	307	16	Q8Y6X1
7	547.5	33.3	307	16	Q92BP4
8	434.5	26.4	299	16	Q8CWM7
9	429.5	26.1	300	16	Q9A1D1
10	428.5	26.1	293	16	Q8P2J5
11	407.5	24.6	298	16	Q8CMP7
12	403.5	24.6	293	16	Q9CHM9
13	396	24.1	300	16	Q8B3T8
14	393	23.9	300	16	Q8DY72
15	392.5	23.9	298	16	Q97PC7
16	217	13.2	343	16	Q8EWK0

17	205.5	12.5	74	2	Q05653	005653 bacillus ce
18	173	10.5	209	2	Q48991	Q48991 mycoplasma
19	168.5	10.3	297	2	Q9RRO3	Q9RRO3 mycoplasma
20	159.5	9.7	259	9	Q9G022	Q9G022 bacterioph
21	151.5	9.2	259	9	Q9B0F8	Q9B0F8 staphylococ
22	146	8.9	257	9	Q9MBR8	Q9MBR8 staphylococ
23	145	8.8	261	16	Q932A3	Q932A3 staphylococ
24	145	8.8	282	16	Q97HS5	Q97HS5 clostridium
25	145	8.8	316	16	Q9PQ14	Q9PQ14 ureaplasma
26	142.5	8.7	296	6	Q38098	Q38098 bacterioph
27	138.5	8.4	235	16	Q67056	Q67056 aquifex ae
28	132	8.0	835	1	Q9V2X2	Q9V2X2 pyrococcus
29	131.5	8.0	329	16	Q97D92	Q97D92 clostridium
30	129	7.9	285	16	Q8E296	Q8E296 leptospira
31	128.5	7.8	739	16	Q51774	Q51774 borrelia bu
32	127	7.7	249	2	Q938D5	Q938D5 escherichia
33	127	7.7	587	17	Q962Y0	Q962Y0 sulfolobus
34	126.5	7.7	327	16	Q8B6U1	Q8B6U1 thermococcus
35	126	7.7	287	16	Q926A4	Q926A4 listeria in
36	121.5	7.4	396	5	Q62556	Q62556 manduca sex
37	120.5	7.3	316	2	P95446	P95446 pseudomonas
38	119	7.2	331	16	Q8KH72	Q8KH72 clostridium
39	118.5	7.2	446	2	Q8GQ84	Q8GQ84 bacillus we
40	118.5	7.2	488	16	Q8RDA3	Q8RDA3 thermococcus
41	117.5	7.2	267	16	Q9K732	Q9K732 bacillus ha
42	116	7.1	842	16	Q98Q49	Q98Q49 mycoplasma
43	114.5	7.0	584	16	Q96Y05	Q96Y05 sulfolobus
44	114	6.9	457	16	Q8XPG2	Q8XPG2 clostridium
45	113.5	6.9	195	17	Q8PTM8	Q8PTM8 methanobarc

ALIGNMENTS

RESULT 1

Q99TH8

PRELIMINARY; PRT; 306 AA.

ID Q99TH8
AC Q99TH8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Primosomal protein.
GN DNAI OR SAV1684 OR SA1507.
OS Staphylococcus aureus (Strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (Strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
[1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus (Strain Mu50), and S. aureus (Strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyma A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kohara S., Goto S., Yabuzaki J.,
RA Kanihisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003363; BAB57846.1; -;
DR EMBL; AP003364; BAB42774.1; -;
KW Complete proteome.
SQ SEQUENCE 306 AA; 35635 MW; 7F3440B89643505E CRC64;

Query Match 97.8%; Score 1607; DB 16; Length 306;
Best Local Similarity 100.0%; Pred. No. 2.4e-115;
Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 8 MGQFSIINTSDPFERIKIKKEVINDDVQKFLFAHABLTNNMIDDLNVLQSYKQ 67
Db 1 MGQFSIINTSDPFERIKIKKEVINDDVQKFLFAHABLTNNMIDDLNVLQSYKQ 60

```
QY 68 QKHVDGHEKADCPNFKVGHVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEELITSHMOR 127
DB 61 QKHVDGHEKADCPNFKVGHVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEELITSHMOR 120
QY 128 DTLNAKLDKIYMNHRDLVAMAAADICTAITNGEQVGLYLPFGTGKSFILGAIANO 187
DB 121 DTLNAKLDKIYMNHRDLVAMAAADICTAITNGEQVGLYLPFGTGKSFILGAIANO 180
QY 188 LKSKKVRSTIIYLPFRITLKGFGKDSFEKKLHVRBANILMLDDIGAEVTPWVDEV 247
DB 181 LKSKKVRSTIIYLPFRITLKGFGKDSFEKKLHVRBANILMLDDIGAEVTPWVDEV 240
QY 248 IGPLHRYMHELPFPSSNFDYSELHHLAMTRDGEKTKAARIIRVKSLSSTPYPLSG 307
DB 241 IGPLHRYMHELPFPSSNFDYSELHHLAMTRDGEKTKAARIIRVKSLSSTPYPLSG 300
QY 308 ENFRNN 313
DB 301 ENFRNN 306
```

RESULT 2

```
Q8NM67 ID Q8NM67 PRELIMINARY; PRT; 306 AA.
AC Q8NM67;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Primosomal protein.
GN DNAI OR MW1627.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=19620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Chai L.,
RA Yamamoto K., Hitamatsu K.,
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827 (2002).
DR EMBL: AP004827; BAB95492.1; -.
KW Complete proteome.
SQ SEQUENCE 306 AA; 35626 MW; CQFCA3752D934B4 CRC64;
```

Query Match 97.3%; Score 1599; DB 16; Length 306;
Best Local Similarity 99.7%; Pred. No. 1e-114;
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 8 MKQPKSIINTSODEPKRIEIKKEVINDDPVKQFLNHRALTNAMIDEDLNLVQEKYDQ 67
DB 1 MKQPKSIINTSODEPKRIEIKKEVINDDPVKQFLNHRALTNAMIDEDLNLVQEKYDQ 60
QY 68 QKHVDGHEKADCPNFKVGHVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEELITSHMOR 127
DB 61 QKHVDGHEKADCPNFKVGHVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEELITSHMOR 120
QY 128 DTLNAKLDKIYMNHRDLVAMAAADICTAITNGEQVGLYLPFGTGKSFILGAIANO 187
DB 121 DTLNAKLDKIYMNHRDLVAMAAADICTAITNGEQVGLYLPFGTGKSFILGAIANO 180
QY 188 LKSKKVRSTIIYLPFRITLKGFGKDSFEKKLHVRBANILMLDDIGAEVTPWVDEV 247
DB 181 LKSKKVRSTIIYLPFRITLKGFGKDSFEKKLHVRBANILMLDDIGAEVTPWVDEV 240
QY 248 IGPLHRYMHELPFPSSNFDYSELHHLAMTRDGEKTKAARIIRVKSLSSTPYPLSG 307
DB 241 IGPLHRYMHELPFPSSNFDYSELHHLAMTRDGEKTKAARIIRVKSLSSTPYPLSG 300
QY 308 ENFRNN 313
DB 301 ENFRNN 306
```

DB 301 ENFRNN 306

RESULT 3

```
Q8CNY2 ID Q8CNY2 PRELIMINARY; PRT; 306 AA.
AC Q8CNY2;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Primosomal protein.
GN SR1358.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 12228;
RA Zhang Y., Ren S., Li H., Fu G., Lu L., Lu G., Jia J., Tu Y., Qin Z.,
RA Chen Z., Wen Y.,
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB016748; AAC04957.1; -.
KW Complete proteome.
SQ SEQUENCE 306 AA; 35736 MW; EDD9F7B52BFC16 CRC64;
```

Query Match 83.7%; Score 1376; DB 16; Length 306;
Best Local Similarity 83.0%; Pred. No. 1.3e-97;
Matches 254; Conservative 33; Mismatches 19; Indels 0; Gaps 0;

```
QY 8 MKQPKSIINTSODEPKRIEIKKEVINDDPVKQFLNHRALTNAMIDEDLNLVQEKYDQ 67
DB 1 MKQPKSIINTSODEPKRIEIKKEVINDDPVKQFLNHRALTNAMIDEDLNLVQEKYDQ 60
QY 68 QKHVDGHEKADCPNFKVGHVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEELITSHMOR 127
DB 61 QKHVDGHEKADCPNFKVGHVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEELITSHMOR 120
QY 128 DTLNAKLDKIYMNHRDLVAMAAADICTAITNGEQVGLYLPFGTGKSFILGAIANO 187
DB 121 DTLNAKLDKIYMNHRDLVAMAAADICTAITNGEQVGLYLPFGTGKSFILGAIANO 180
QY 188 LKSKKVRSTIIYLPFRITLKGFGKDSFEKKLHVRBANILMLDDIGAEVTPWVDEV 247
DB 181 LKSKKVRSTIIYLPFRITLKGFGKDSFEKKLHVRBANILMLDDIGAEVTPWVDEV 240
QY 248 IGPLHRYMHELPFPSSNFDYSELHHLAMTRDGEKTKAARIIRVKSLSSTPYPLSG 307
DB 241 IGPLHRYMHELPFPSSNFDYSELHHLAMTRDGEKTKAARIIRVKSLSSTPYPLSG 300
QY 308 ENFRNN 313
DB 301 ENFRNN 306
```

RESULT 4

```
Q8CXB0 ID Q8CXB0 PRELIMINARY; PRT; 308 AA.
AC Q8CXB0;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Primosome component (Helicase loader).
GN DNAI OR OB2156.
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Oceanobacillus.
NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=2220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.,
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
RT Ridge and its unexpected adaptive capabilities to extreme
```

RT environments. 30:3927-3935 (2002).
 RL Nucleic Acids Res. 30:3927-3935 (2002).
 DR EMBL: AP004600; BAC14112.1; -
 KW Helicase; Complete proteome.
 SQ SEQUENCE 308 AA; 35729 MW; 8D4CF70B480A833 CRC64;

Query Match 37.8%; Score 621.5; DB 16; Length 308;
 Best Local Similarity 41.4%; Pred. No. 9.5e-40;
 Matches 127; Conservative 67; Mismatches 110; Indels 3; Gaps 3;

QY 7 IMQFSSINTSDPFERIKIKIKVINDPVKQFLFAHRAELTNAMIDEDLVNLOEYKD 66
 DB 4 ISELKAMKKNKPFQTSYQNIKREKVIDPSIKHFL-ANHEBELSEIIDKHLITLYEYKS 62
 QY 67 QOHHYDGHK-FADCPNVKGVHPELVYDNNRIKIRYLCQPCPKIKYDERFEALITSHM 125
 DB 63 QSHQCDCKSLGGQNNIQQSPVLEADNNDIRISTYKCHRLSEEGNEQOQKLIQSLYM 122
 QY 126 QBDTLNAKLDIYNNHRDLVDVMAADDICTALTNGEYKGLYVPGFGKSPILGATA 185
 DB 123 PKHILQARISDVQDEH-RSNALGKVLDFLEASKEQLPKKGLYVGSFGVGTLYLGATA 181
 QY 186 NOLKSKKVRSTIYLPFIRTLKGFQDSPEKKLHVRREANILMDIGAEVTPWVRD 245
 DB 182 NELKLEYSISLYMPEFVEIKSSFQDSFNEKVDFFKADITIMLDGAEWQSAWFRD 241
 QY 246 EVLGPILHYRMVHELPFFSSNPDYSELEHHLMTGRDEKTKYARIRIEVKSISTYFL 305
 DB 242 EVLASVLYQYMMGELPFVITSNDYDLQELSTTRNGEVQVAGRILIERIKOYTTDKL 301
 QY 306 SGENFRN 312
 DB 302 SGENFRS 308

RESULT 5
 QY9K863 PRELIMINARY; PRT; 311 AA.

AC Q9K863;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Primosome component (Helicase loader).
 GN DNAI OR BH3144.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL: AP001517; BAB06863.1; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR SMART: SM00382; AAA; 1.
 KM Helicase; Complete proteome.
 SQ SEQUENCE 311 AA; 35822 MW; 6AFB3073431363BB CRC64;

Query Match 36.6%; Score 601.5; DB 16; Length 311;
 Best Local Similarity 39.2%; Pred. No. 3.3e-38;
 Matches 125; Conservative 69; Mismatches 104; Indels 21; Gaps 8;

QY 5 OSIMKQKSLINSDPEKRIKIKVINDPVKQFLFAHRAELTNAMIDEDLVNLOEY 64
 DB 2 ESIQSSIKHMTG-NRSEKQGTOLKEAVFRSPHVLLEER-DELSPTTLEQGSKEYEY 59
 QY 65 KDOQKHYDGH-KFADCPNVKGVHPELVYDNNRIKIRYLCQPCPKIKYDERFEALIT 122

DB 60 QKEQSHC-AHCPGLQKCPNLMKGYQPTLYVERDSELSYSPCLKEEERKKRSLIRS 118
 QY 123 HHMQBRTLVNAKLDI-----YNNHRDLVDVMAADDICTALTNGEYKGLYVPGFGT 175
 DB 119 LVPKEILEKFPDVSEBGRSTASHR-ALEPLSKP-----SDGKGLYVGFV 170
 QY 176 GKSPILGATANOLSKSKVSTIYLPFIRTLKGFQDSPEKKLHVRREANILMDIG 235
 DB 171 GKTFLMGATANELKONGIDSTIYVDFPRELKQSIGDGFQOQKLPVNAOVLITPDDIG 230
 QY 236 ABEFTVWVDEVIQPLHYRMVHELPFFSSNPDYSELEHHLMT-RDGEKTKYARIRIE 294
 DB 231 AETMTSVWVDVGVILQYRIMEKLPFTLSNDYDELSEHLAYNDSGTLLKAKVME 290
 QY 295 RVKSLSTPYFLSGENFRN 313
 DB 291 RINHYTSVWVQGNRRH 309

RESULT 6

ID Q8Y6X1 PRELIMINARY; PRT; 307 AA.
 AC Q8Y6X1;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Primosome component (Helicase loader) Dnai.
 GN DNAI OR LMO1560.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=BGD-e / Serovar 1/2a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glasier P., Frangoul L., Buchrieser C., Ruehnik C., Amend A.,
 RA Baquero F., Berche P., Bioecker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetaniani F., Couve E., de Daruvar A., Deloux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fehli H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Gobeil W., Gomez-Lopez N., Hain T., Haul J., Jackson D.,
 RA Jones L.-M., Kaerle U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 RA Meduno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
 RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tietze A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species.";
 RL Science 294:849-852(2001).
 DR EMBL: AL591979; CAC99638.1; -
 DR Listliet; LMO01560; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR SMART: SM00382; AAA; 1.
 KM Helicase; Complete proteome.
 SQ SEQUENCE 307 AA; 35970 MW; CF2FD9350CEB0F48 CRC64;

Query Match 33.6%; Score 552.5; DB 16; Length 307;
 Best Local Similarity 37.8%; Pred. No. 1.9e-34;
 Matches 113; Conservative 69; Mismatches 104; Indels 13; Gaps 4;

QY 19 QDEKRIKIKVINDPVKQFLFAHRAELTNAMIDEDLVNLOEYKDOQKHYDGHK 78
 DB 15 RDEKRYQGLKQGVHAYQPIQDPFKHKEVTEQLVYNOLSNLYEWTGHKKTBE--- 71
 QY 79 CPNFVKGHVELYDNNRIKIRYLCQPCPKIKYDERFEALITSHMQBRTLVNAKLDIY 138
 DB 72 -ETLMGVA PKVLVNEEFITVYYPKEKIEEDKRAVRRRIISLTPMPQVADANADFY 130
 QY 139 MNHRDLVDVMAADDICTALTN-----GEOYGLYVPGFGKSPILGATANOLSKKY 193
 DB 131 TDESR---QLALVEYQPLANNPPKSGRVKGLFTHSGTGKSTYLGALAEALMKG 187
 QY 194 RSTIYLPFIRTLKGFQDSPEKKLHVRREANILMDIGAEVTPWVRDEVIGPLH 253

```

Db      186 STTVLVLPFPMKMEVQKQISDNTVGEKIQGAKSTETVLMDDIGAESMTAMTNRBVGAILQ 24
Qy      254 YRMVHELPTFFSSNDFDYSLEHNLAMTRDG-EKTKAARIIRVXKSLSTPYTLGSENR 311
          :::::::::::::::::::::
Db      248 FRMQELPTFFSSNNYMQLEHNLHMFACNGTEKTKARIMERVYLSKRVNLGKNR 306

RESULT 7
ID      092BF4      PRELIMINARY,      PRT,      307 AA.
AC      092BF4;
DT      01-DEC-2001 (TREMBlrel. 19, Created)
DR      01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE      01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE      Prionosome component (Helicase loader) DnaI.
GN      DnaI OR L1N1595.
OS      Listeria innocua.
OC      Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX      NCBI_TaxId=1642;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CLIP 11262 / Serovar 6a;
RX      PubMed=11679669;
RA      Glaeser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA      Baquero F., Berche P., Bloeker H., Brandt P., Chattraporty T.,
RA      Chabdic A., Chetouiati F., Couve E., de Darvar A., Deloux P.,
RA      Domnan B., Dominguez-Bernal G., Duchaud E., Durant L., Dussauget O.,
RA      Eutlian K.-D., Fsihi H., Garcia-del Portillo F., Garido P.,
RA      Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA      Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunze F., Kurupkhat G.,
RA      Madeno E., Maitouram A., Mata Vicente J., Ng E.-Z., Nedjati H.,
RA      Nordstak G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA      Remmel B., Rose M., Schueter T., Simoes N., Tlertaz A.,
RA      Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.;
RT      "Comparative genomes of Listeria species.";
RL      Science 294:849-852(2001)
DR      EMBL, AL596163; CAC96826.1; -.
DR      L1ec11a1c; L1N01595; -.
DR      InterPro; IPR003593; AAA_ATPase.
DR      SMART; SMO0382; AAA; 1.
KW      Helicase; Complete proteome.
SQ      SEQUENCE 307 AA; 35988 MW; A8517663BA8F123F CRC64;

Query Match      33.3%; Score 547.5; DB 16; Length 307;
Best Local Similarity 37.8%; Pred. No. 4.5e-34;
Matches 112; Conservative 68; Mismatches 109; Indels 7; Gaps 3;

Qy      19 ODFEKRIKKEVINDPVGQFLAHRAELTNAMIDEDLVLYQKQCKHYDGHKFPAD 78
          :::::::::::::::::::::
Db      15 KDFEKYVGLKQOVLAHYQPIODFFENKEDITELQINOLSLYFWMTQHKKFTQOE--- 71

Qy      79 CPNFVKGVPELVYVNNRIKIRYLOCPCKIKYDEERFEAEALITSHMORDTLNAKDIY 138
          :::::::::::::::::::::
Db      72 -ETLMGVAAPKVLNNGEFTIYVYPTKEIIEBDRRAVBRRIKSLYMPKQVVDANLAFY 130

Qy      139 MNHRDRLVMAAADICTAI--TNGEQVKGVLVYPPFGTGSFLLIGALANOLSKKXVST 196
          :::::::::::::::::::::
Db      131 TDEBSRKALVAYEAYOFLNINYPQNEERYVGLFIHSGFGKSYLLIGALAKELALKGISTY 190

Qy      197 IIVPEFIRTLKGGKQDSFEKFLHRAVNAITMLDIIAGEVTPWVWDEVIGLLHVRM 256
          :::::::::::::::::::::
Db      191 LVYLPEFMEKVEQGISDNTVGEKIQPAKETEVLMLDDIGAESMTAMTNRBDEVIGAILQFRM 250

Qy      257 VHELPTFFSSNDFDYSLEHNLAMTRDG-EKTKAARIIRVXKSLSTPYTLGSENR 311
          :::::::::::::::::::::
Db      251 QBELPTFFSSNNYMQLEHNLHMFACNGTEKTKARIMERVYLSKRVNLGKNR 306

RESULT 8
ID      08CWM7      PRELIMINARY,      PRT,      299 AA.
AC      08CWM7;
DT      01-MAR-2003 (TREMBlrel. 23, Created)

```

```

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative DNA replication protein, primosome component (Helicase
DE loader).
DE DNAI OR SMU.1921.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
NCBI_TaxID=1309;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Mcshan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson W.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najaf F., Lai H., White J., Roe B.A., Ferretti J.J.,
RA "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RA pathogen."
RT Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439 (2002).
RW EMBL; AB015016; AAN59532.1; -.
DR Helicase; Complete proteome.
SQ SEQUENCE 299 AA; 34410 MW; 98CAAB56F5524BD1 CRC64;

Query Match 26.4%; Score 434.5; DB 16; Length 299;
Best Local Similarity 32.0%; Pred. No. 2e-25;
Matches 99; Conservative 68; Mismatches 127; Indels 15; Gaps 5

QY 4 GOSIMKQPKSIINTSODPEKRIEIKKEVINDDPVKQFLAHRARLITNMIDEDLNVLOE 63
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 5 GQTLAKDNHKNYDT-----EKMTQALISDQEIANTIAH--HLSGEQKISLPKFNQ 54
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 64 YKDOQKHYDGKFPADCPNFVGNVPELYVDNNRKIRLQCPCKIKDDEEFAELITSH 123
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 55 YRLRRTREPMH---DQAVYIAKGYOPLVVMNCGYADVAKETKEELIAAKKTAISPRINVV 111
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 124 HMOBDPLTNAKLKDIYMHNRDLVDVAMADDICTAITNGEYVKGSLYLYGPFQKSFILGA 183
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 112 SLRPSYKNISPDIDNLVDVKFLDVPKRVADVEQTPNEQ--KGLYLYGDMGIGSYLMAA 170
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 184 IANOLSKKK-VRSTIIYLPERIRTLKGGFGKDSFEKKLHVRANITLMLDDIGAEVTPW 242
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 171 MAHELSSEGRGAATTLHPSPSTIDVKNAINGTVAKEIDAVKTDILIDDIGAEQSTSW 230
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 243 VRDEYIGPLHYRNVHLELPFPSSNPFDSSELEHHLNMTROSEBKTAKARIIRKVSISTP 302
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 231 IRDELQVILQYRLMLELPLFPFTSNYSFKDLKANKLNKGSDETWQAKRVMERIRYLAKE 290
   ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 303 YFLSGENFR 311
   |||||
Db 291 IHLSEGENFR 299

RESULT 9
09A1D1 PRELIMINARY; PRT; 300 AA.
AC 09A1D1
AC 09A1D1
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Putative primosome component (Helicase loader).
DE DNAI OR SFY0340 OR SPYK3_0248.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus
NCBI_TaxID=1314, 198466;
RN (1)
RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=SF370 / ATCC 700294 / serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., Mcshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najaf F.Z., Ran Q., Zhu H., Song L., White J.,

```

RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.pyogenes; STRAIN=MGAS315 / Serotype M3;
 RX MEDLINE=22133808; PubMed=2122206;
 RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
 RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
 RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
 RA Schlievert P.M., Musser J.M.,
 RT Genome sequence of a serotype M3 strain of group A *Streptococcus*:
 RT phage-encoded toxins, the high-virulence phenotype, and clone
 RT emergence."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
 DR EMBL: AEO06498; AAK33392.1; -
 DR EMBL: AEO14141; AAM78855.1; -
 KW Hypothetical protein; Helicase; Complete proteome.
 SQ SEQUENCE 300 AA; 34146 MW; 489827A789338C5A CRC64;
 Query Match 26.1%; Score 429.5; DB 16; Length 300;
 Best Local Similarity 31.2%; Pred. No. 4.9e-25;
 Matches 94; Conservative 74; Mismatches 126; Indels 7; Gaps 4;
 QY 12 KSIIINSDPEKRIEIKKEKVINPDVQKFLAHRALTNAMIDEDLVLOEKDOCKY 71
 DB 6 ETMAKLGQNTNRVNSDOLITLADPEVASISOH-HLSOQINLSKNOFLVEROKY 63
 QY 72 DGHKPADCPNFVKGHVELVYNNRIKIRYLQCPCKIKYDEERFEALITSHMQDPTLN 131
 DB 64 ---QLNDPSYIAGYOPILAMNEGADVSYLTKELVEAQKQAIISRIQLVELPKSYRH 120
 QY 132 AKLKQIVMHRDLVDAMADDICTAITNGEYKGLYLPFGTSGSFILGAIANOLSKR 191
 DB 121 IHLSDIDVNNASWEPASALIDFVEQYPSAQ-KGLYLVGDWIGISYLLAAVAHLSK 179
 QY 192 K-VRSTIIYLPFIRTLKGFQKDFEKKLHRYREANIIMLDIGABEVTWVRDEVIGP 250
 DB 180 KGVSTLLHFPSPALIDVKNASVKEIDAVKPNVILDDIGABQATSWRDEVLYQ 239
 QY 251 LHYRMVHELPTFPSSNFDVSELEHHLAMTRDESEKTKAARIIEKYSITPFLSGENP 310
 DB 240 ILQYRLMELPTFPSTSNVSPADLERKWKATIKGSDETWQAKRVVERVYLAERFHLGANKR 299
 QY 311 R 311
 DB 300 R 300
 RESULT 10
 Q88235 PRELIMINARY; PRT; 293 AA.
 AC Q8P25;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Putative primosome component (helicase loader).
 GN SPM18_0332.
 OS *Streptococcus pyogenes* (serotype M18).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxId=186103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MGAS8232 / serotype M18;
 RX MEDLINE=21927593; PubMed=11917108;
 RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
 RA Sylva G.L., Sturdevant D.B., Ricklefs S.M., Porcella S.F.,
 RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
 RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.,
 RT "Genome sequence and comparative microarray analysis of serotype M18
 RT group A *Streptococcus* strains associated with acute rheumatic fever
 RT outbreaks."

RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
 DR EMBL: AEO09978; AAL97087.1; -
 KW Helicase; Complete proteome; Hypothetical protein.
 SQ SEQUENCE 293 AA; 33357 MW; 0E87AD94BDB599B6 CRC64;
 Query Match 26.1%; Score 428.5; DB 16; Length 293;
 Best Local Similarity 32.4%; Pred. No. 5.7e-25;
 Matches 93; Conservative 70; Mismatches 117; Indels 7; Gaps 4;
 QY 26 ERIKKEVINPDVQKFLAHRALTNAMIDEDLVLOEKDOCKYHDKFKADCPNVKG 85
 DB 13 DDLITLADPEVASISOH-HLSOQINLSKNOFLVEROKY---QLNDPSYIAG 67
 QY 86 HVELVYNNRIKIRYLQCPCKIKYDEERFEALITSHMQDPTLNAKLKDITMNRDL 145
 DB 68 YQPIAMNEGADVSYLTKELVEAQKQAIISRIQLVELPKSYRHILSDIDVNNASVK 127
 QY 146 DVAMADDICTAITNGEYKGLYLPFGTSGSFILGAIANOLSKK-VRSTIIYLPF 204
 DB 128 EAFSALIDFVEQYPSAQ-KGLYLVGDWIGISYLLAAVAHLSKGVSTLLHFPFA 186
 QY 205 RTLKGFQKDFEKKLHRYREANIIMLDIGABEVTWVRDEVIGPLHYRMVHELPTFP 264
 DB 187 IDVKNAISNGSVKEIDAVKPNVILDDIGABQATSWRDEVLYQLQYRLMELPTFP 246
 QY 265 SGNPDVSELEHHLAMTRDESEKTKAARIIEKYSITPFLSGENP 311
 DB 247 TSNVSPADLERKWKATIKGSDETWQAKRVVERVYLAERFHLGANKR 293
 RESULT 11
 Q88CMP7 PRELIMINARY; PRT; 298 AA.
 AC Q8CMP7;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Primosome component (Helicase loader).
 GN DNAT OR SPM155.
 OS *Streptococcus pneumoniae* (strain ATCC BAA-255 / R6).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 NCBI_TaxId=171101;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429245; PubMed=11544234;
 RA Hoskins J., Alborn W.B. Jr., Arnold J., Blaszcak L.C., Burgess S.,
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
 RA Gilmore R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsumura P.,
 RA Mahlen S.M., McHenry M., McLeaster K., Mundy C.W., Nicas P.I.,
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rokey P.,
 RA Sun P.-M., Winkler M.B., Yang Y., Young-Bellido M., Zhao Q.,
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
 RA Glass J.I.,
 RT "Genome of the bacterium *Streptococcus pneumoniae* strain R6."
 RL J. Bacteriol. 183:5709-5717(2001).
 DR EMBL: AEO08523; AAL00359.1; -
 KW Helicase; Complete proteome.
 SQ SEQUENCE 298 AA; 34574 MW; 3EA8DBD1DEB8F22 CRC64;
 Query Match 24.8%; Score 407.5; DB 16; Length 298;
 Best Local Similarity 32.0%; Pred. No. 2.4e-23;
 Matches 94; Conservative 64; Mismatches 121; Indels 15; Gaps 7;
 QY 21 FKRRIKKEVINPDVQKFLAHRALTNAMIDEDLVLOEK-KDOCKYDGHKPADCP 79
 DB 17 YQDLVQKIMK---DPDVAATVQ--QESLNGDELNRSSKFMQYITERKFLRG----DT 66
 QY 80 PNFVKGHVELVYNNRIKIRYLQCPCKIKYDEERFEALITSHMQDPTLNAKLKDITM 139
 DB 67 DYIAGYKRIWVNHGADVSYETPTELLAARKEAIAKRLNLINPSSLKVVSPADVYR 126

RA DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R., Lewis M.R.,
 RA Radu D., Fedorova N.B., Scanlan D., Kouri H., Mulligan S.,
 RA Cary H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
 RA Jacobini E.T., Brettoni C., Gall G., Mariani M., Vogni F., Malone D.,
 RA Rimando D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative genomic analysis of an
 RT emerging human pathogen, serotype V Streptococcus agalactiae";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396(2002).
 DR EMBL: AB014265; AAN00485.1; -.
 DR TIGR: SAG1621; -.
 KW Complete proteome.
 SQ SEQUENCE 300 AA; 34663 MW; 9BED36F86B363B9 CRC64;
 Query Match 23.9%; Score 393; DB 16; Length 300;
 Best Local Similarity 28.5%; Pred. No. 3.1e-22;
 Matches 88; Conservative 80; Mismatches 127; Indels 14; Gaps 5;
 QY 4 GOSIMQFQSIINTSODFEKRIKKEKVINDDVAKQFLAHPAELTNAMIDEDLVLOE 63
 DB 5 GQALENGQVPRNTN-----DELQIMILADQVAFETIT--QLSQREINISMKEFNQ 55
 QY 64 YKDQKHVDGKFKADCPNFKVGVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEILTSH 123
 DB 56 FLIERQKF---KKKDSQYIAKGYEPILVMNEGADVSYLETRELIIRAQKQAIISDRINLV 112
 QY 124 HMDRDLNKLKDIYNNHRLVMAAADICTAINTGEBOVKGLIYGPFGTKSFILGA 183
 DB 113 NLPKSYRNIRMTDFDINNBSRMKMSQLDFEVTYSYNN-KELYLYGDMGVKSYLMAA 171
 QY 184 IANOLSKK-VRESTIYLPFIRTLKGGFKDGSFEKKLHVRBANIMLMDIGAEVTPW 242
 DB 172 MARELSERKQVSTTLHFPFALDVKAISSTGYKDEIDAVKSVPIILDDIGAEQATSW 231
 QY 243 VRDEVIGPLHYMVHELPFPSSNPDYSLEHHLAMTRDGEKTKAARIIERYKSLSTP 302
 DB 232 VRDEILQVILQHRMLEELPFPFTSNYSFNDLEKRMANIKGSDTQAKRVMERVRYLAIE 291
 QY 303 YPLSGENR 311
 DB 292 FPLEGNRR 300
 RESULT 15
 Q97PC7 PRELIMINARY; PRT; 298 AA.
 AC Q97PC7, 01-OCT-2001 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Primosomal protein DnaI.
 GN SP1711.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TIGR4;
 RX MEDLINE=21357209; PubMed=1463916;
 RA Tettelin H., Nelson K.B., Paulsen I.T., Bisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radu D.,
 RA Holtzapfel E., Kouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angluoi S., Dickinson T., Hickey E.K.,
 RA Holt I.B., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae";
 RL Science 293:498-506(2001).
 DR EMBL: AB007464; AAK75789.1; -.
 DR TIGR: SP1711; -.

KW Complete proteome.
 SQ SEQUENCE 298 AA; 34259 MW; CAE5ECAFA9C30059 CRC64;
 Query Match 23.9%; Score 392.5; DB 16; Length 298;
 Best Local Similarity 32.3%; Pred. No. 3.3e-22;
 Matches 96; Conservative 64; Mismatches 116; Indels 21; Gaps 8;
 QY 21 FEKRIEIKKEVINDDVAKQFLAHPAELTNAMIDEDLVLOEY-KDQKHVDGKFKADC 79
 DB 17 YQDLVQIKK---DDVNAFIO--QESLTPKELNRISIKFQYTERKFLRG----DT 66
 QY 80 PNFVKGVPELVYDNNRIKIRYLOCPCKIKYDEERFEAEILTSHMDRDLNKLKDIYM 139
 DB 67 DYIAGYKPIVMNNGVADVSEETBELIAEKKAIRKRLKLNIPASLKKASIAQVVL 126
 QY 140 NHRDLV---AMADDICTAINTGEBOVKGLIYGPFGTKSFILGAINOLSKK-VRS 195
 DB 127 DDLGRLPVFEKLIAPVEQYPAIR-----KGLIYGPFGVGSFMAALAHDLSEKRGVSS 181
 QY 196 TIYLPFIRTLKGGFKDGSFEKKLHVRBANIMLMDIGAEVTPWVRDEVIGPLHYR 255
 DB 182 TLHLPSPFVIDVKMAISDGNVKTLYDEIKLSVLLDDIGAEOSTVWVRDEVILQVILQYR 241
 QY 256 MVHELPFPSSNPDYSLEHHLAMTRDGEKTKAARIIERYKSLSTPYPLSGENR 311
 DB 242 MQENLPFTSNFPNFEDEKHFPAKVHGNDETWEARRVWERIRYLAETRLLEGVNR 298
 Search completed: December 15, 2003, 15:15:56
 Job time : 39.1651 secs

THIS PAGE BLANK (USPTO)